

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 10, 2004, 08:11:50 ; Search time 4658 Seconds  
(without alignments)  
3024.149 Million cell updates/sec

Title: US-09-787-737-2  
Perfect score: 1751  
Sequence: 1 MSSSNKWPMSFKSPKNNN.....TDFGVGYQLONGAIYILI 325

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DRV=xlh  
-O=/cgn2\_1/USPTO\_spool\_h/US09787737/runat\_01092004\_112454\_23595/app\_query.fasta\_1.519  
-Db=GenEmbl -OEW=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09787737 @CGN 1 1 3731 @runat\_01092004\_112454\_23595 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DAV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1751	100.0	1214	6	E52177	B52177 Homeobox ge
2	1751	100.0	1214	6	BD012130	BD012130 Homeobox
3	1586	90.6	84510	8	AB006698	AB006698 Arabidops
4	645.5	36.9	1518	6	E52178	B52178 Homeobox ge
5	645.5	36.9	1518	6	BD012131	BD012131 Homeobox
6	643.5	36.8	1500	8	AK118501	AK118501 Arabidops
7	576	32.9	114144	8	U78721	U78721 Arabidopsis
8	420.5	24.0	918	8	BT005296	BT005296 Arabidops
9	394.5	22.5	807	8	BT006232	BT006232 Arabidops
10	394.5	22.5	996	8	AK117910	AK117910 Arabidops
11	369.5	21.1	159878	2	AC124143	AC124143 Oryza sat
12	369.5	21.1	182868	2	AC121360	AC121360 Oryza sat
13	347.5	19.8	182681	8	AP003760	AP003760 Oryza sat
14	338	19.3	1282	8	AK063503	AK063503 Oryza sat
15	338	19.3	1378	8	AK073232	AK073232 Oryza sat
16	337.5	19.3	81701	8	AB006706	AB006706 Arabidops
17	328.5	18.8	1367	6	AX105291	AX105291 Sequence
18	312	17.8	1449	8	AK111000	AK111000 Oryza sat
19	312	17.8	1602	8	AK063262	AK063262 Oryza sat
20	307.5	17.6	89334	8	ATAC009327	ATAC009327 Arabidops
21	299.5	17.1	1247	8	AK107305	AK107305 Oryza sat
22	280.5	16.0	67448	8	AP005249	AP005249 Oryza sat
23	237.5	13.6	152365	8	AP005167	AP005167 Oryza sat
24	236	13.5	162711	8	AC137070	AC137070 Genomic s
25	213	12.2	103125	8	AC005560	AC005560 Arabidops
26	212.5	12.1	1326	8	AY085832	AY085832 Arabidops
27	209.5	12.0	1117	8	AY162209	AY162209 Antirrhin
28	200	11.4	1209	8	AY086385	AY086385 Arabidops
29	199.5	11.4	1227	8	ATH012310	ATH012310 Arabidops
30	194.5	11.1	1137	8	AF481951	AF481951 Petunia x
31	193	11.0	88149	8	AC006201	AC006201 Arabidops
32	192.5	11.0	131246	8	CNS08CDT	CNS080503 Oryza sat
33	192.5	11.0	147501	2	CNS09584R	EX072548 Oryza sat
34	192.5	11.0	161808	8	CNS08CDM	EX000496 Oryza sat
35	192	11.0	1180	8	AK102203	AK102203 Oryza sat
36	189.5	10.8	76444	8	AP006099	AP006099 Lotus cor
37	187.5	10.7	1338	6	AX105285	AX105285 Sequence
38	187	10.7	741	6	AX105301	AX105301 Sequence
39	187	10.7	1035	6	LES538329	AJ538329 Lycopersi
40	187	10.7	14305	2	AC137078	AC137078 Arabidops
41	186.5	10.7	74968	8	AB019230	AB019230 Arabidops
42	183	10.5	134982	8	AP003279	AP003279 Oryza sat
43	183	10.5	150976	8	AP003442	AP003442 Oryza sat
44	178.5	10.2	749	8	AY150812	AY150812 Arabidops
45	178.5	10.2	74494	8	AC007171	AC007171 Arabidops

ALIGNMENTS

E52177	E52177	Homeobox gene encoding protein participating in differentiation.	1214 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	DEFINITION					
E52177	ACCESSION					
E52177.1	VERSION	GI:18629640				
JP 2001029081-A/1.	KEYWORDS	Arabidopsis thaliana (thale cress)				
Arabidopsis thaliana	ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.						
1 (bases 1 to 1214)						
Kakimoto, T.	REFERENCE					
Homeobox gene encoding protein participating in differentiation	AUTHORS					
Patent: JP 2001029081-A 1 06-FEB-2001;	TITLE					
SUNTORY LTD, NIPPON PAPER IND CO LTD	JOURNAL					
OS Arabidopsis thaliana (thale cress)						
PN JP 2001029081-A/1	COMMENT					
PD 06-FEB-2001						
PF 22-JUL-1999 JP 1999207995						
PI TATSUO KAKIMOTO						
PC C12N15/09, A01H1/00, A01H5/00, C07K14/42, C12N1/15, C12N1/19, PC						
C12N1/21, C12N5/10,						
PC C12P21/02, C12N15/00, C12N5/00						
CC						
Key Location/Qualifiers						
FT CDS						
1..1214						
/organism="Arabidopsis thaliana"						
/mol_type="genomic DNA"						
/db_xref="taxon:3702"						
ORIGIN						
Alignment Scores:						
Pred. No.:	1..11e-157	Length:	1214			
Score:	1751.00	Matches:	325			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			
US-09-787-737-2 (1-325) x E52177 (1-1214)						
QY 1 MetSerSerAsnLysAsnTyrProSerMetPheLysSerLysProCysAsnAsn 20						
DB 36 ATGTCCTCCCAACAAATAATGGCCAGCATGTTCAATCCAAACCTTGCAACAATAAT 95						
QY 21 HisHisGlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40						
DB 96 CATCATCATCAACATGAATCGATACCTCATCTTACATGCACTACTCTAATTCGAACCTA 155						
QY 41 SerSerSerPheSerAspArgLysProAspProLysProAspTyrAsnProLysPro 60						
DB 156 TCATCTCTCTTCTCCAGATCGGATACCATCTCAATCCAGATGGAATCCTTAACCG 215						
QY 61 GluGlnLeuArgGluLeuSerPheAsnSerGlyThrLeuAsnProProArgGlu 80						
DB 216 GAGCAGATTAGATACCTCAATCAATCTTCAATCCGGTACTATTAAACCCACCTAGAG 275						
QY 81 GluLeuGlnArgGluLeuArgLeuGlnGluTyrGlyGlnLeuGlyAspAlaAsnVal 100						
DB 276 GAGATTCAAGAATCCGATCCGGCTTCAAGAATATGTCCTCAATCCGGTACCAACCGT 335						
QY 101 PheTyrTyrPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120						
DB 336 TTTTACTGGTTTCAAAACCGGAATCTCGAGCAAAACCAACAGCTTCGTTCATCACAA 395						
QY 121 SerProLysMetSerLysLysAspLysThrValLeuProSerThrAspAlaAspHisCys 140						
DB 396 AGCCCTAAATGTCAAAGAAGACAAAGACGGTTATTCTCTAGTACTGACGCTGATCATGTT 455						
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160						
DB 456 TTTGGTTTGTTAACCAAGAACCCGATATATCCGGTTCAAAACAATGAGTTGGTGA 515						
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaGlnSerAla 180						
DB 516 ACCGAACCGCGCGTTTCTATTCCGGTTCAATATGATCCGAGCGCTCTCAATCAGCG 575						
QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200						
DB 576 TTTGGTTTGGCGATTTTGTGTACCGGTGGTAACGGAAGAGGAGGATGCAATCTCTACC 635						
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysLeuProAlaLeuAsn 220						
DB 636 GTTAATAACGGGTTAATTTGGAGACTAACGAAATTTTGTATAAAATTCGGCGATCAAT 695						
QY 221 LeuTyrGlyGlyAspGlyAsnGlyGlyLysCysPheProProLeuThrValProLeu 240						
DB 696 TTATACGGCGGAGATCGAAATGCGGTGGAATTTGTTTCTCTCTTGAATGTTCCATTA 755						
QY 241 ThrLeuAsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyGluAspValGly 260						
DB 756 ACCATCAATCAATCTCAAGAAACAGAGATGATGATTATCCGTTGTGTAGACGTCCGA 815						
QY 261 AspAsnValTyrProValArgMetThrValPheLeuAsnGluMetProLysGluValVal 280						
DB 816 GATTAATGTTTATCCGGTGAGATGACGGTGTATTATTAACGAGATGCTATCGAAGTAGTG 875						
QY 281 SerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValLeuLeuAsnSerPhe 300						
DB 876 TCTGGATTATTCAACGTTAAGCGAGCTTTTCGGAACGATCCGCTTTTGATCAACTCGTTT 935						
QY 301 GlyGlnProLysLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyVal 320						
DB 936 GGCCAGCCTATTCTTACGATGAATTTGGTGTACTTATCACTTATCAACTCTCCAAATGGCGCA 995						
QY 321 IleTyrTyrLeuLeu 325						
DB 996 ATCTATTATCTTATT 1010						
RESULT 2						
BD012130	LOCUS	Homeobox gene coding for protein participating in differentiation.	1214 bp	DNA	linear	PAT 02-AUG-2002
DEFINITION						
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
COMMENT						
OS Arabidopsis thaliana (thale cress)						
PN WO 0107618-A/1						
PD 01-FEB-2001						
PF 21-JUL-2000 WO 2000JP004904						
PR 22-JUL-1999 JP 99P 207995						
PI TATSUO KAKIMOTO						
PC C12N15/29, C12N1/21, C07K14/415, C12P21/02, C12N5/10, A01H5/00// PC						
(C12N15/29, C12R1/91), (C12N1/21, C12R1/01), (C12P21/02, C12R1/01) CC						
Nucleotide sequence coding for protein participating in differentiation						
CC						
Key Location/Qualifiers						
FT CDS						
1..1214						
/organism="Arabidopsis thaliana"						
FEATURES						
source						

/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"

## ORIGIN

Alignment Scores: Pred. No.: 1,11e-157 Length: 1214  
Score: 1751.00 Matches: 325  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-787-737-2 (1-325) x BD012130 (1-1214)

QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsnAsn 20  
DB 36 ATGTCCTCCTCAACAAAATGGCCAGCATGTTCAAATCCAAACCTTGCAACAATAAT 95  
QY 21 HisHisGlnHisGluIleAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40  
DB 96 CATCATCATCAACATGAATCGATACCTCCATCTTACATGCACACTCTAATTGCAACCTA 155  
QY 41 SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60  
DB 156 TCATCTTCCTTTTCCTCAGATCGGATACCAAGATCCTAAACCGAGATCGTAACCG 215  
QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGlu 80  
DB 216 GAGCAGATTAGGATACGATCGAATCAATCTTCAATTCGGTACTATTAAACCCACCTAGAG 275  
QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100  
DB 276 GAGATTCAAGAATCCGGATCCGGCTTCAAGATATGTCATCAATCGTGACGCAACCGTG 335  
QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120  
DB 336 TTTTACTGGTTTCAAAACCGGAATCTCGAGCAAAACACAAAGCTTCGTGTTCATCAAAA 395  
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAsnHisCys 140  
DB 396 AGCCCTAAATGTCAAGAGGACAGACGGTTATTCTAGTACTGACGCTGATCATTTGT 455  
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160  
DB 456 TTTTGGTTTGTGTTAAACCAAGAACCGGATTATATCCGGTTCAAAACCAATGATGGTGGTA 515  
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaGlnSerAla 180  
DB 516 ACCGAACCGCGCGGTTTCTATTTCGGTTCATAATGATCCGAGCGCTGCTCAATCAGCG 575  
QY 181 PheGlyPheGlyAspPheValProValProValThrGluGluGlyMetAlaPheSerThr 200  
DB 576 TTTTGGTTTGGCGATTGTTGTTACCGGTGGTAAACGGAAGAGGATGCGATCTCTAC 635  
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsn 220  
DB 636 GTTAATAACGGGTGTAATTTGGAGACTAACGAAAAATTTTGATAAAATTCGCGCATCAT 695  
QY 221 LeuTyrGlyGlyAspGlyAsnGlyGlyAsnCysPheProProLeuThrValProLeu 240  
DB 696 TTATACGGCGGAGATGGAATGGCGGTGGAATTTGTTTCCCTCTTCACTGTTCAATTA 755  
QY 241 ThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyLysValGly 260  
DB 756 ACCATCATCATCTCAAGAAAAACGAGATGAGGATTATCCGGTGGTGAAGACGTCGGA 815  
QY 261 AspAsnValTyrProValArgMetThrValPheIleAsnGluMetProIleGluValVal 280  
DB 816 GATAATGTTTATCCGGTGAGAAATGACGGTGTGTTTATTACGAGATCGCTATCGAAGTAGT 875  
QY 281 SerGlyLeuPheAsnValLysAlaPheGlyAsnAspAlaValLeuIleAsnSerPhe 300  
DB 876 TCTGGATTATTCAAGGTTAAGGCAGCTTTCGGAACGATCGCGTTTTCATCAACTCGTTT 935

QY 301 GlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAla 320  
DB 936 GGCCAGCCTATTCTTACAGATGAATTTGGTGTACTTATCAACCTCTCCAAATGGCGCA 995  
QY 321 IleTyrTyrLeuIle 325  
DB 996 ATCTATTATCTTATT 1010

## RESULT 3

AB006698

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

PUBMED

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

LOCUS

DEFINITION

AB006698 BA000015

ACCESSION

AB006698.1 GI:2351063

VERSION

KEYWORDS

SOURCE

ORGANISM

/protein\_id="BAB08242.1"  
/db\_xref="GI:9757717"  
/translation="MSTARFKICVTVGAVGKTCMLISYTSNTPTDYVPTVDFNFS  
ANVVDGSGVNLGMDTAGQBYNLRPLSYRGADVFLLAFLSLIGKASYENHKKWLP  
ELKHVAPGIPVILVGTKLDRDKQKLDKHPGASITTAQGBELKMGIGAVRYLECSS  
KIQQNVKAVFDFAIRVALRPKAKKKIKLTKRSKICEFL"  
/note="contains similarity to homeodomain transcription  
factor  
gene\_id:MCL19.2"  
/codon\_start=1  
/evidence=not\_experimental  
/product="CRN (crooked neck) protein"  
/protein\_id="BAB08244.1"  
/db\_xref="GI:9757718"  
/translation="MSSSNKNMPSMFKSKPCNNHHQHHEIDTPSMYHNSCNLSFF  
SDRIIPDPKRWPKPEQRIILIESFNSGTINPPREIIRIRIQEYQIGDANVFF  
WQNRKRAKHLRVHHSKPKSKDKTVLPSTADHCFGVNQETGLYPVQNNELVV  
TPAGFLFVHNDPSAAQAFGRGDFVVPVTEEGVAFSTVNVNLETFNENFDKIPA  
INLYGDGNGGNCFFPLIPVITINOSQKLLISK"  
/note="10120,10206. .10271,10355. .10499,10600. .10818,  
join(9000. .10120,10206. .10271,10355. .10499,10600. .10818,  
10909. .11071,11182. .11377,11457. .11568)  
/notes="gene\_id:MCL19.3"  
/codon\_start=1  
/evidence=not\_experimental  
/product="CRN (crooked neck) protein"  
/protein\_id="BAB08244.1"  
/db\_xref="GI:9757719"  
/translations="MTGDAEVLKPRTRVKNKTPAPVQITARQILREAREQAEIR  
PPKQKTDHLSLDRRRKEPEDQIRRARWNIQVVKVAKVEESQMDYARASVME  
RALEGYRHTLWVKAPEPMKNVNNARVNDRSVTLPLPRVDOLWEKVIYVKEKLG  
NVGAKQIFERNWNSPDQKAWLCFKIPELRYNEIRARSIIYERVLCHKVSAFIRY  
AKFEMRGQVKLAREYERAVDKLANDEAEILVFSFAEFERKEVERKARTIYKPA  
LDHIRKRAEELYKFAVEKQYQKEGIEDAIVGKRFEYEDVSKNPLNYSWFDY  
VLYEESGKNDRIREYERAIANVPAQEKRFQRYTYLWINTALYEEIEKDVERT  
DRVECKLIPHTKPSFAKTLAAEVEIRQLMTGROILGNAIGKAPVKVIFKKYI  
EMELKLVNDRCKLIERELFENSPENYANVAREPISLAETERARAPELAI SOPA  
LMPPELLWIKITIDFELISEGFETKRALYERLLORTKCKVWISFAKPEASAKHGD  
IKSARVIFRANTYKIDTTPLEBEERATLLEDMLNMTGEGELGSDVSVOGSKLPKLLK  
KRWTSREDSGEYERFDYLPFEESGTNNKILEAAKWKLLKGEVCV"  
/note="join(11995. .12297,12529. .12809,12975. .13130,  
13269. .13404,13629. .13736,13846. .14052)  
/notes="gene\_id:MCL19.4"  
/codon\_start=1  
/evidence=not\_experimental  
/product="myrosinase-binding protein-like; jasmonate  
inducible protein-like"  
/protein\_id="BAB08245.1"  
/db\_xref="GI:9757720"  
/translation="MSKVIYAGSGNGKFDGQVYSGTVGVKLVGDEHGIYVL  
KIQYKNGDVVVEHGRARGTHITETFEVKCPDEYITISWGTGRNHHRYKSGRT  
FEQITADKYDINLAARMGTPMYNTVSELOFTKSHGRTSEQFRMPTGGSGVSWDD  
GAYDRLSKLCVGEDAHCVSSEYFHYVKGNDRIHCHGKDSKEHDKGFISSLTKTSM  
NRSEKFKGPVGFQLEAKGDFKIVGFRGKSVNRIALGANFAVAVVPPVKLINAK  
GGVLGNEWDIGLHDDAMITFKLYFNKEYITISVGHYKRLAAPNASASMSPTGY  
MTMLKFTNTRITVQLSHSPEYTEGTSFKLEEDKHLVGVGYKTESLQIGVYVXP  
IANA"  
/note="18260. .18398,18630. .18859)  
/notes="gene\_id:MCL19.5"  
/unknown protein"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAB08246.1"  
/db\_xref="GI:9757721"  
/translation="MSSSNKNKNKDDHDDHNDQGHNTHTPSFMHTFESNTSSPS  
SPVYSDPKPEKPNQHOAQILEELFIGTWNPSLTSIKQITIKLSQSGEYEDDADVYK  
WFHNRKYSRKEPLYSFMFL"  
/note="join(20751. .20785,20889. .20954,21047. .21132,  
21235. .21269,21369. .21455,21539. .21616,21701. .21741,  
22239. .22305)  
/notes="gene\_id:MCL19.6"  
/unknown protein"  
/codon\_start=1  
/evidence=not\_experimental

/protein\_id="BAB08247.1"  
/db\_xref="GI:9757722"  
/translation="MGRCKFKKPTQORFSSAADILAGTSAAPRSPFKQKAEYBED  
VEEESSEESDEADYKKGAEAVIEVDPNVRQKTLKAKLDASKTTELSPRE  
REELRQAHARYMELQOGKTEQARKDLRLALLRQREEAARKREKAARDAKKV  
EGRK"  
/complement(join(22934. .23259,23634. .23851))  
/notes="gene\_id:MCL19.7"  
/pseudo  
/codon\_start=1  
/evidence=not\_experimental  
/product="Rab-type small GTP-binding protein-like"  
/join(27158. .27278,27480. .27606,27840. .27915,28196. .28338,  
28394. .28464,28552. .28573,28658. .28814)  
/notes="gene\_id:MCL19.8  
pir||T27174  
/similar to unknown protein"  
/codon\_start=1  
/evidence=not\_experimental  
/protein\_id="BAB08248.1"  
/db\_xref="GI:9757723"  
/translation="MLGHSSMDNYASDEVTSWAFLLLLAFSTRTVSPDCLEREIRI  
KTMKREVERNEEKALWFRQGFNFDLVKCLILSKLNGHLHLSPPORDHRLLENL  
NNTHPLVLKQWERGNOESLLRSEWISLTQSLVLSAITGLVSNASKMHNIDMKH  
LICKAACRICESFSTITALTEAIDIYSEWIDECERVANTAEDDVVQEEEEVEE  
EEREDEDDHYSVKRYNF"  
/complement(join(29073. .29934,30037. .30602,30797. .31014,  
31113. .31227))  
/notes="gene\_id:MCL19.9"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08249.1"  
/db\_xref="GI:9757724"  
/translation="MTVEEVGDDYTDGTVDLRGNVRRSQTGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSSVSQITKQGRGVIQNNLSRDLNYYMFAVLNLLNLLFLVIRFVYRAD  
VTOSANVQKEPNVDNYNE"  
/complement(join(32738. .33188,33279. .33677,33771. .34336,  
35792. .36009,36633. .36747))  
/notes="gene\_id:MCL19.10"  
/codon\_start=1  
/evidence=not\_experimental  
/product="peptide transporter"  
/protein\_id="BAB08250.1"  
/db\_xref="GI:9757725"  
/translation="MTVEEVGDDYTDGTVDLQCNVRRSIRGRWKACSFVVYVEVE  
RMAYGISNLIYMTTKLHQGTVKSSNNVTNWSLTPILGAVVADAHGRITTF  
VTSAYLGLMALLTSLVSLGPKPCSTANVENCKASVIGLAVFEGALYLAIGT  
GGPKNTITGADQDFEDPKDKIKHHSFNNWMSIFPGTFATVTVVQDNVGA  
IGYCLTGLAFSIFLLGTRLYRHKLPMGSPPTKMARVIVASLRKAREPSSDSTR  
FYGLPMEYAKARAFIHSSTSLRPLNRASLKTGSHKWLCTITEVEETKQMLMLP  
VLETFVPSMMLAQIMTLFIKQTLDRRLTNFSPAPSLGFTTSMVLSVLYDR  
PIPLSPTLLPOVLMGLADAFIEIAKLEFFYDQAPESMKSGLSTSTSNAGVFW  
VYVFKRKLTPNGRGTILQRMGIMLHILIMIAISITERKLVKAAEGHGHQTA  
SILLSS

[illegible]

[illegible]

```
DB: 6 Gaps: 17
US-09-787-737-2 (1-325) x BD012131 (1-1518)
QY 1 MetSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsn 20
152 ATGGCTTCTTTCGAATAGACATCGCCAGCATGTTCAAGTCCAAACCT-----CATCC 205
QY 21 HishHisGlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40
206 CATCAATGGCAACATGACATCACTCTCT-----CTCTGGCTTCTGCTTCACCGA 259
QY 41 SerSerSerPheSerSer-----AspArgIleProAspProLysProArgTTP 56
260 TCTTCTCTCTCTCTCAGGATGTGAAGTGAGAGGAGTCCAGAGCCAAACCAAGATGG 319
QY 57 AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsn 76
320 AATCCAAAGCCAGACGATTCGGATATCTTGAAGCAATCTTTAACTCCGGGATGGTGAAT 379
QY 77 ProProArgGluGluIleGlnArgIleArgLeuGlnGluTyrGlyGlnIleGly 96
380 CCTCCAGAGAGGAGATCAGAGGATVAGGCTCAGCTTCAAGATACGGCCAAAGTCGGT 439
QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
440 GATGCTAAACGCTTCTACTGTTCCAAACCGTAAGTCCGCTAGTAAACACAACTCCGC 499
QY 117 ValHisHis----- 119
D5 500 CTCTCCCAACCACTCCAAACACTCTCTCCCTCAAACGCAACCGCAGCCGCGCAA 559
QY 120 -----LysSerProLysMetSerLysLysAspLysThrValIlePro--- 133
560 CCTTCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 619
QY 134 -----SerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGly 149
620 AAAAGCAAGAACAAAGAACAACTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 676
QY 150 LeuTyrProValGlnAsnGluLeuValThrGluProAlaGlyPheLeuPhePro 169
677 ATGTTTCCA-----CCGGAACCGGCG-----TTCTCTCTCTCTCTCTCTCTCT 709
QY 170 ValHisAsn-----AspProSerAlaAlaGlnSerAlaPheGlyPhe----- 183
710 GTCCTCACTCGGAGGGTTGAAGTATACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
QY 184 GlyAspPheVal-----ValProValValThrGluGluGlyMetAlaPheSerThr 200
770 GGTGATATGATTGAGCAACAAAAACCGGCTCCAAACGTGTACCGGACTCTCTGCTGAGT 829
QY 201 ValAsnAsnGly---ValAsnLeuGluThrAsnGlu----- 211
830 ATCATGACGGTAGTGTGATGATTATGGAAGTATCATCAACACACTTCAGTGAGAAAGAA 889
QY 211 ----- 211
890 GTTGAAGAAATCAGGATGAGATGTTCACAGCCACAGACTCAGATTTGTTCACCTACC 949
QY 212 AsnPheAspLysIleProAlaIleAsnLeuTyrGlyAspGlyAsnGlyGlyAsn 231
950 ACTAATCATCAATAGCTCTCTACAAAC-----AACAAACACACACACATTAAC 997
QY 232 Cys-----PheProPro----- 235
998 ATCATGCTTCATATTCCTCCCACTACTTCTACTGCCACCACTATTACTACTTCGCATTCT 1057
QY 236 ---LeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer 254
1058 CTCGCTACTGTCCCATCACTTCGACACCAAGCTTCAAGTTCAGGCGGAC----- 1105
QY 255 GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu 274
```

```
1106 -----GCACGATAAGAGTATTTTCATCAATGAA 1132
QY 275 MetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAla 294
1133 ATGGAGCTTGAAGTGAAGTCTAGGACCCCTTCAATGTGAGGATGTCATTTGGGGAAGAGTT 1192
QY 295 ValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGln 314
1193 GTTCTGATTAAATCCCGGGGTGAGCCCATTTGTCACCATGATGAATATGGCGTCTCTTCAC 1252
QY 315 ProteuGlnAsnGlyAlaIleTyrTyrIleuIle 325
1253 CTTCTTCAACAGGAGCCTCGTACTATCTGATC 1285

RESULT 6
AKI18501 1500 bp mRNA linear PLN 06-DEC-2002
LOCUS Arabidopsis thaliana At2g33880 mRNA for putative homeodomain
DEFINITION transcription factor, complete cds, clone: RAFL19-73-K21.
ACCESSION AKI18501
VERSION AKI18501.1 GI:26452029
KEYWORDS FLI cDNA; CAP trapper
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
Kawai,J., Hayashizaki,Y. and Shinozaki,K.
JOURNAL Arabidopsis thaliana full-length cDNA
REFERENCE 2 (bases 1 to 1500)
AUTHORS Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
Hayashizaki,Y. and Shinozaki,K.
JOURNAL Direct Submission
COMMENT Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Suehiro-cho Tsurumi-ku, Yokohama, Kanagawa
230-0045, Japan (E-mail:mseki@gsc.riken.go.jp,
URL:http://pfgweb.gsc.riken.go.jp, Tel:81-45-503-9625,
Fax:81-45-503-9586)
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720;
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified lambda FLC-1-E vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
This clone is in a modified pluescript vector.
Please visit our web site (http://pfgweb.gsc.riken.go.jp/) for
further details.
FEATURES
Location/Qualifiers
1..1500
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAFL19-73-K21"
/note="common name: thale cress"
1..1500
/gene="At2g33880"
380..1297
/gene="At2g33880"
/codon_start=1
/product="putative homeodomain transcription factor"
/protein_id="BAC43105.1"
/db_xref="GI:26452030"
translation="MVNPPREIRIRIAQLQEVGVGDANFYHFQNRKSKHKLRL
LNHSKSLPQTQPOQPQPSASSSSSSSSSSSKTKPRKKNNTNLSIGGQMMG
MFPPEAPLFPVSTVGEGITVSSQLGFLSGMIRQKPAKPTCTGLLSIMNGSVS
YGHQHQHLSKEVEEMRMKWLQOPQTQICVYATNHQIATSYNNNNNNNNIIMLHPPTT
```



STATTTTSHSLATVPSTSDQLQVQADARIRVFINEMELELVSSGPFNVDRDAFGBEVL  
INSAGQPIVTDYGVALHPLOHGASYILI"

## ORIGIN

Alignment Scores:  
Pred. No.: 8, 94e-52 Length: 1500  
Score: 643.50 Matches: 166  
Percent Similarity: 50.61% Conservatives: 42  
Best Local Similarity: 40.39% Mismatches: 84  
Query Match: 36.75% Indels: 119  
DB: 17 Gaps: 17

US-09-787-737-2 (1-325) x AK118501 (1-1500)

QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsn 20  
DB 161 ATGGCTTCTTGAATAGACACTGGCCAAAGCATGTTCAAGTCCAAACCT-----CATCCC 214  
QY 21 HisHisHisGlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40  
DB 215 CATCAATGACACATACATCACTCTCCT-----CTCTGGCTTCTGCTTCTCACCGA 268  
QY 41 SerSerSerPheSerSer-----AspArgIleProAspProLysProArgTrp 56  
DB 269 TCTTCTCCTTCTCTTCCAGATGTGAAGTGGAGGAGTCCAGAGCCAAACCAAGATGG 328  
QY 57 AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsn 76  
DB 329 AATCCAAAGCAGAGCAGATTCGATACCTTGAAGCAATCTTAACTCCGGGATGTGAAT 388  
QY 77 ProProArgGluGluLeuGlnArgIleArgIleArgLeuGlnGlnIleGly 96  
DB 389 CCTCCAAAGAGAGAGATCAGAGGATTAGGGCTCAGCTTCAAGAAATACGCCAAGTCGGT 448  
QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116  
DB 449 GATGCTAAAGCTCTTACTGTTCCAAAACCGTAAGTCCCGTAGTAAACACAACTCCGC 508  
QY- 117 ValHisHis----- 119  
DB 509 CTCCTCCACACCACTCCAAACACTCTCTCCCTCAAAACGCAACCGCAGCCGCGCAA 568  
QY 120 -----LysSerProLysMetSerLysLysAspLysThrValIlePro--- 133  
DB 569 CTTTCGGCTTCT 628  
QY 134 -----SerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGly 149  
DB 629 AAAAGCAAGACAAAGACACACTAAATCTCTCTTTGGT---GGTAGTCAAATGATGGG 685  
QY 150 LeuTyrProValGlnAsnAsnGluLeuValThrGluProAlaGlyPheLeuPhePro 169  
DB 686 ATGTTTCCA-----CCGGAACCGGG-----TTTCTCTTCCCG 718  
QY 170 ValHisAsn-----AspProSerAlaAlaGlnSerAlaPheGlyPhe----- 183  
DB 719 GTCTCCACTGTCGGAGGTTTGAAGGTATCACCCTCTCAUCCCAATAGGGTTTCTCTCC 778  
QY 184 GlyAspPheVal-----ValProValValThrGluGluGlyMetAlaPheSerThr 200  
DB 779 GGTGATGATGATGACAAACAAACCGGCTCAACAGTGTACCGGACTCCTGCTGAGTGAG 838  
QY 201 ValAsnAsnGly-----ValAsnLeuGluThrAsnGlu----- 211  
DB 839 ATCATGAACGGTAGTGTGATGTTATGGAACTCATCATCAACACACTTGGTGAAGAA 898  
QY 211 ----- 211  
DB 899 GTTGAAGAAATCAGGATGAAGATGTGCAACAGCCACAGACTCAGATTGTTAGCTACC 958  
QY 212 AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGlyAsn 231  
DB 959 ACTAATCATCAATAGTCTTCTTACAC-----AACCAACAAACAAACAAATAAC 1006

QY 232 Cys-----PheProPro----- 235  
DB 1007 ATCATGCTTCATATTCTCTCCCACTACTTCTACTGCCACACTATTACTACTTCGCATTCT 1066  
QY 236 ---LeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer 254  
DB 1067 CTCGGTACTGTCTCCATCACTTCGGACCACTTCAAGTTCAAGCGGAC----- 1114  
QY 255 GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu 274  
DB 1115 -----GCAGGATAGAGTTCATCATGAA 1141  
QY 275 MetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAla 294  
DB 1142 ATGGAGCTTGAAGTGAAGCTCAGCGCTTCAATGTGAGGATGCATTCGGGAAGAGTT 1201  
QY 295 ValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGln 314  
DB 1202 GTTCTGATTAATTCGGGGTCCAGCCATTGCACCGATGAATATGGCGTTCGCTCTTCC 1261  
QY 315 ProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325  
DB 1262 CCTCTTCAACAGCGGACCTCGTACTATCTGATC 1294  
RESULT 7  
U78721 114144 bp DNA linear PLN 27-FEB-2002  
LOCUS Arabidopsis thaliana chromosome 2 clone T1B8 map TEN5, complete  
DEFINITION sequence.  
ACCESSION U78721.3 GI:20198308  
VERSION U78721  
KEYWORDS HTG.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 114144)  
AUTHORS Rounsley, S.D., Lin, X., Ketchum, K.A., Phillips, C.A., Brandon, R.C.,  
Fuhrmann, J.L., White, O., Kerlavage, A.R., Adams, M.D.,  
Somerville, C.R. and Venter, J.C.  
UNPUBLISHED  
REFERENCE 2 (bases 1 to 114144)  
JOURNAL Lin.X.  
AUTHORS Direct Submission  
TITLE Submitted (09-MAR-2000) The Institute for Genomic Research, 9712  
JOURNAL Medical Center Dr., Rockville, MD 20850, USA  
REFERENCE 3 (bases 1 to 114144)  
JOURNAL Town, C.D. and Kaul, S.  
AUTHORS Direct Submission  
TITLE Submitted (27-FEB-2002) The Institute for Genomic Research, 9712  
JOURNAL Medical Center Dr., Rockville, MD 20850, USA, cdtonet.igb.org  
COMMENT On Apr 18, 2002 this sequence version replaced gi:6598813.  
FEATURES  
Location/Qualifiers  
1..114144  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/map="TEN5"  
/clone="T1B8"  
1..1789  
/note="overlap with BAC clone F4P9  
(AC002332:150805..152593)."  
misc\_feature  
1..931  
/gene="At2g33670"  
/note="synonym: T1B8.33"  
join(1..153,246..281,358..425,534..594,692..931)  
/gene="At2g33670"  
join(1..153,246..281,358..425,534..594,692..931)  
/gene="At2g33670"  
misc\_feature  
1..1789  
/note="overlap with BAC clone F4P9  
(AC002332:150805..152593)."  
1..931  
/gene="At2g33670"  
/note="synonym: T1B8.33"  
join(1..153,246..281,358..425,534..594,692..931)  
/gene="At2g33670"  
join(1..153,246..281,358..425,534..594,692..931)  
/gene="At2g33670"



```

/note="Protein sequence is in conflict with the conceptual
translation"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69142.3"
/db_xref="GI:20198313"
/translation="KQAIISDMWALEIOERHAYIQGMPLNVNSDRHFWRPALVLHI
IHILFNQAEITYFFWYFGLRSCFHHFALIIIRVALGVGVQLCSYITLPLVA
LVQMGTMRKSVFDDTSKALKNWKNAKKSTPGQTQPLPLNLPKTKGGIESAS
PANITASVDVKESDQSQSRDLLSGP"
gene
2660.4843
/gene="At2g33680"
/note="synonym: TlB8.1; predicted by genscan and
genefinder, similar to GP 1402883 gnl PID e248504 X98130"
<2660..>4843
/gene="At2g33680"
2660.4843
/gene="At2g33680"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69139.1"
/db_xref="GI:1707007"
/translation="MHPSFQTELNPHSTLLKKLTHSQORNLVAGRAVHGQIIIRT
ASNCIGHANVLNYPYAKGKIAKAHSIFNAIIKOVVSNLSITGYSONGGISSYTV
MQFLREARQDILPNAYTLAGIFKAESSQSSTVGRQAAHLYVMSFGDIYVDTSLV
GMVYKAGLVDGLKVFAYPERNTYTWSTWVSGYATGRVBAIKVFNLFLEKRGSG
DSDYFTAVLSLAATYVGLGROIHCITITKNGLLGFVALSNALVTMYSKCESINEAC
KNPDSGDRNSITWSAMVTGYSONGSELAVKLFSPMSAGIKPSTYITVGVNLNACSD
ICVLEGRKOLHSLKLGPERHLFATLVDNPTAKGLADARKDPDCIOERDVALWT
SILSGVQNSDNEBALIYRMTKTAGIIPNDPTMASVLKASLATLELGKQVHGHTI
KIGFLEVPISGLPMYSKCSLEGNLFRTPNKDVVSNAMI SGLSHNGQDEA
LELFEMLAEGEPDPTVNTIISASHKGFVRGWFYFNWMSDQIGLDPKVDHYACM
VOLLRAGOLKAEKERTEGANIDHGLKWLRIILLSACKNHGKCELGIVYAGEKLMALGSR
ESSTYVQLSGITVIALGRMDVEFVKHMRANGSVKSGVCSWELKNQIYFVVGDTMH
PMLEETKDACLVSROMIEGFVTLDSSFVEEBEGTQLSTSFII"
5251..5466
/gene="At2g33690"
/note="synonym: TlB8.27; predicted by genscan and
genefinder"
<5251..>5466
/gene="At2g33690"
5251.5466
/gene="At2g33690"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69143.1"
/db_xref="GI:2253578"
/translation="MSKSEKQELPLETSPTYKYEDIEDYKKNAYGTSGHQDVKPGHG
GGTTDPTFGDAAPSAIDSANQAKK"
complement(6459..6492)
7252..9045
/gene="At2g33700"
/note="synonym: TlB8.2"
join(<7252..7434,7522..7713,7864..8069,8275..9045)
/gene="At2g33700"
join(7252..7434,7522..7713,7864..8069,8275..8836)
/gene="At2g33700"
/codon_start=1
/product="putative protein phosphatase 2C"
/protein_id="AAC69126.1"
/db_xref="GI:1707015"
/translation="WMSMDFSPLLTVLEGDNKNTSSATETDLENLDDTQISKGRP
PHLTSSATRLQLAANDVDVNCNLMKSLDDKSPFLPYRSGCAEQAKQFMDDEHI
CIDILVNHLGAIQCSSIGARYGYFDGAGTDAHFVRKNILIRFIVEDSSPFLCVKKA
IKSAPLKADYFADDSLIDSSGTGHTAFIFGRRLITANAGDCRAVLGRGRAIELS
KDHPCNTAEKREKLGVVYDGLQSVARAIGDWHMKPGKACPLSPPELQ
ETDLSDEDFLIMGDCGLWDMVMSQCVTIARKELIMIHNDPERCSRELREALKNTIC
DNLATVIVCFPPDPQRIEIRMQSRVRSISAEGLNLLKGVLDGYP"
complement(11621..11649)
/rpt_family="AT rich"
/rpt_family="AT rich"
/gene="At2g33710"

/note="synonym: TlB8.3"
complement(join(<12323..12798,13524..>13704))
/gene="At2g33710"
complement(join(12323..12798,13524..13704))
/gene="At2g33710"
/codon_start=1
/product="putative AP2 domain transcription factor"
/protein_id="AAC69127.1"
/db_xref="GI:1707016"
/translation="MHSGKRLPESMAGNREEKKEIACCSTLSESDVSFVSELTQQ
PIPSDDOSSLTLOKSNRORNYGRQRPWGKWAERIDPNKARARVWLGTFTDA
EEAALAVDKAAFERGHKAKLNPEHIRVNPTOLYFSPATSHDRILIVPPSPPTIAP
DILLDQGHGFQSRSSSSANLSNMWLSSSSSSLHQLGRLENEDGENVKNISIKRRK"
complement(14124..14156)
/rpt_family="AT rich"
complement(14385..14425)
/rpt_family="(TA)n"
16828..18041
/gene="At2g33720"
/note="synonym: TlB8.29; predicted by genscan and
genefinder"
join(<16828..17139,17373..>18041)
/gene="At2g33720"
join(16828..17139,17373..18041)
/gene="At2g33720"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69145.1"
/db_xref="GI:2253580"
/translation="MKMPFFFSASKTYQLFHDSESSKSLVSTLSSCSBNPKR
KNSDEVNISCIPRYDKLTQVERKIARNLSYQKAEDEWGVSTELTLPKDPWII
KKVFFASVLDMPNVSNTLCLDTESENAESVLSLDLCFYDKNTWPHDNVANK
PTESEAINLWMTSKRKEEKYVSTELTLLTADPWLTKWMTKSSGNLRL
VLKSAFVDIHLRLPLDDQMWKEDSGLAVHYDHTDTSVENLAKWKSSSSFLV
SGWRKCFVDRGLQGVDTGVTWDRSESKLHFVLSRSETMSAPLPSP"
complement(18731..20932)
/gene="At2g33730"
/note="synonym: TlB8.4"
complement(<18731..>20932)
/gene="At2g33730"
complement(18731..20932)
/gene="At2g33730"
/codon_start=1
/product="putative U5 small nuclear ribonucleoprotein, an
RNA helicase"
/protein_id="AAC69128.1"
/db_xref="GI:1707017"
/translation="MKRTINEVVGSTVNTLDSSIAKPVFLTKKORELAKRRQDQ
ISQRVREQLGRPEPETEDVSGDNRDKDRDRDRDRDRDRDRDRDRDRDRDR
DRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDR
DAI KEQYLGKPKKRVIRPEKFPFSFDWENTEDTSRDMNVLQNPHQAQLFQRF
RHMDSKLEEMTERDWRIPREDFNISYKSGRIPRPMRSEESKLTSELLKAVERAGY
KPSIQMAAIFLQQRDVIAGTSGKTAFLVPLMAYISRLPPMSENETEGFYA
VMAPTRELAQIIEETVFAHYLGFRTVTSVGGQSEEOGLKTQGCIVATPGRL
IDCLERYAVLNQNYVVLDEADMIDMPFPQVAGVLDAMPSSNLKPENEERLDEK
KIYRTYMFSAIMPFGVNRKARLKNPVVVTICTAGTKTDLISCHVIMKESKFFR
LOKLLDELGEKTAIVFVNTKNCDSIAKLDKAGRVVTLHGKKSQOEIISLGFRA
KRYNVLVATDVVVGIGDIPVAHVINDMPKHIEMTHYTHIRIGTRAGKSGVATSLTL
HDTBFVLDLKLQVNSAVPELPALARHSEARFKFTGTPDRPRHSHTVYIN"
complement(20588..20718)
/rpt_family="(CAGAG)n"
complement(21321..22493)
/note="synonym: TlB8.17; supported by full length cdna:
Ceres: 269036"
complement(join(<21321..21655,21744..21814,22051..22126,
22221..22298,22384..>22493))
/gene="At2g33735"
complement(join(21584..21655,21744..21814,22051..22126,
22221..22298,22384..22446))
/gene="At2g33735"

```

```

/codon_start=1
/product="Expressed protein"
/protein_id="AAM15518.1"
/db_xref="GI:20198310"
/translation="MWCSWDGSHGFDFEDYKDYKLVLELNCDSDEIRSFIR
LALKHPDKFKEDSATSRFQELNEAYQVLSPIARQYDKRMRIYEHNTLELNEY
KELIITCNGLGKVKRYLW"
repeat_region 22524..22544

Alignment Scores:
Pred. No.: 6 56e-43 Length: 114144
Score: 576.00 Matches: 166
Percent Similarity: 40.53% Conservatives: 48
Best Local Similarity: 31.44% Mismatches: 93
Query Match: 32.90% Indels: 221
DB: Gaps: 17

US-09-787-737-2 (1-325) x U78721 (1-114144)
QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsnAsn 20
Db 96652 ATGGCTTCTTCGAATAGACACTGCCAAGCATGTTCAAGTCCAAACCT-----CATCC 96599
QY 21 HisHisHisGlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40
Db 96598 CATCAATGGCAACATGACATCAACTCTCTCT-----CTCTGCTTCTGCTTCTCACC 96545
QY 41 SerSerSerPheSer----- 45
Db 96544 TCTTCTCTCTCTCTCTCAGGTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 96485
QY 46 -----Ser-Aspar 48
Db 96484 TTGTTTTAAAGAACTCTTATCTGATGAGATGTGAATTTTTTTAGGATGTGAAGTGAGAG 96425
QY 48 gileProAspProLysProLysProLysProGluGlnLeuArgLeuGluSe 68
Db 96424 GAGTCCAGAGCCAAACCAAGATGGAATCCAAAGCCAGACAGATTCGGATCTTGAAGC 96365
QY 68 rlePheAsnSerGlyThrLeuAsnProProArgGluGluLeuArgLeuArgLeuAr 88
Db 96364 AATCTTAACTCCGGATGTTGATATCTCCAGAGAGAGATCAGGAGGATTAGGGCTCA 96305
QY 88 gLeuGlnGluTyrGlyGlnLeuGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgL 108
Db 96304 GCTTCAAGAAATACGCCAAGTCGGTGTGCTAAGCTTCTTACTGTTTCCAAACCCGTAA 96245
QY 108 sSerArgAlaLysHisLysLysLeuArgValHisHis----- 119
Db 96244 GTCCCGTAGTAAACACAAACTCCGCTCTCTCCACCAACCACTCTCTCTCTCTCTCTCT 96185
QY 120 -----LysSerProLysMetSerLy 126
Db 96184 AACCAACCGCAGCGCAGCGCAACCTTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 96125
QY 126 sLysAspLysThrValLePro-----SerThrAspAlaAspHisCysPh 141
Db 96124 CTCCTCCAAATCCCAACCAACCCGAAAGAAAGCAAGCAACCACTAATCTCTCTCT 96065
QY 141 eGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValValTh 161
Db 96064 GGGT---GGTAGTCAATGATGGGATGTTTCCA-----CC 96032
QY 161 rGluProAlaGlyPheLeuPheProValHisAsn-----AspProSerAlaAl 177
Db 96031 GGAACCGCGG---TTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 95975
QY 177 aGlnSerAlaPheGlyPhe-----GlyAspPheVal-----ValProValValTh 192
Db 95974 CTCATCCCAATTAGGCTTCTCTCCGGGATGATGATGAGCAACAAACCGGCTCCAC 95915
QY 192 rGluGluGlyMetAlaPheSerThrValAsnAsnGly---ValAsnLeuGluThrAsnGl 211

```

```

Db 95914 GTGTACCGGACTCCTGCTGAGTGAGATCATGAACGGTAGTGAGTTATGGAACTCATCA 95855
QY 211 u----- 211
Db 95854 TCAACAACACTGAGTGAGAAAGAGTTGAAGAAATGAGGATGAAGATGTTGCAACAGCC 95795
QY 212 -----AsnPheAspLysIleProAlaIleAsnLeuTyrGl 223
Db 95794 ACAGACTCAGATTTCTGAGCTTACCACTAATCATCAAAATAGCTTCTTCAAC- 95743
QY 223 yGlyAspGlyAsnGlyGlyGlyAsnCys-----PheProPro----- 235
Db 95742 -----AACACACACAAACAATAACATCATCATGCTTCTCTCTCTCTCTCTCTCTCTCT 95687
QY 236 -----LeuThrValProLeuThrIleAsnGlnSerGl 246
Db 95686 CACCACTATTACTACTCTCGCATCTCTCGCTACTTCCCATCACTTCGGACCGCTTCA 95627
QY 246 nGlyLysArgAspValGly----- 252
Db 95626 AGTTCAGAGTAATGTTAATATATATAAAGTTATCACCCTTTTAAAGTTATAGATTACTTATA 95567
QY 252 ----- 252
Db 95566 TACATGTTGAAATTTTACTCAATCGTAGAGATACATTAATTTTGTGTCAACATGTAAG 95507
QY 252 ----- 252
Db 95506 GACACATTCATCCATACAAATCTTTTAAACCTTTCATGTTTAGTGTAGCTTTAGATGA 95447
QY 252 ----- 252
Db 95446 TTCTTTATTTTATATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 95387
QY 253 -----LeuSerGlyGlyGluAs 258
Db 95386 TTTTGTTTTATAGATTGTTATAGTATATATATACATGAAATATATGGAATAACAATG 95327
QY 258 pValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGluMetProIleGl 278
Db 95326 GCTAATATGATGCGAGCGGACGACGACGATTAAGATTTTTCATCAATGAAATGAGCTTGA 95267
QY 278 uValValSerGlyLeuPheAsnValIleAlaPheGlyAsnAspAlaValLeuIleAs 298
Db 95266 AGTGAGCTCAGGACCGTTCATGAGGATGATGATGAGGATGATGAGGATGATGAGGATG 95207
QY 298 nSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAs 318
Db 95206 TTCGGGGTCCAGCCCATGTCCCGATGATGATGATGATGATGATGATGATGATGATGAT 95147
QY 318 nGlyAlaIleTyrTyrLeuIle 325
Db 95146 CGGAGCTCTGACTATCTGCTT 95125

```

RESULT 8  
BT005296 918 bp mRNA linear PLN 14-MAR-2003  
LOCUS Arabidopsis thaliana At2g33880 mRNA, complete cds.  
DEFINITION Arabidopsis thaliana (thale cress)  
ACCESSION BT005296  
VERSION BT005296.1 GI:28950872  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
1 (bases 1 to 918)  
Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,  
Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,Y.W.,  
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,  
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,  
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Seki,M.,  
Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,

**TITLE**  
JOURNAL  
**REFERENCE**  
**AUTHORS**

Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
Arabidopsis ORF clones  
2 (bases 1 to 918)  
Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P., Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuai,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

**TITLE**  
JOURNAL

**COMMENT**

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : "RIKEN Arabidopsis Full-length cDNA") : Seki,M., Narusaka,M., Ishida,J., Sacou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki.Y. and Shinozaki.K.  
The Salk, Stanford, PGRC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Chan, M.M., Chang, C.M., Dale,J.M., Hsuai,V.W., Jones.T., Karlin-Neumann.G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm.C.J., Quach,H.L., Southwick.A., Tang,C.C., Toriumi.M., Wong,C., Wu.H.C., Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and Ecker,J.R.

**FEATURES**  
source

Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.  
Location/Qualifiers  
1..918  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U60909"  
/ecotype="Columbia"  
/note="This clone is in pUNI 51"  
1..918  
/note="putative homeodomain transcription factor"  
/codon\_start=1  
/product="At2g33880"  
/protein\_id="A2Q63360.1"  
/db\_xref="GI:28950873"  
/translation="WVNPPEIRRIRAQLQYGVGDANVFYWFQNKRKSRSHKLRL LHNKHSLUPQPQQPPQSASSSSSSSKTKPKRSKNNTNLISLGGSGMML YPFPPFAFLFYTVVGFGIGTIVSQGLSDGMIEQOKPATCTCGLLILSETMGVS VSPHQAHLSEKEVMEMRMKLQQPTOICYATNHOIASYNNNNNNNIIMLPPTT STATTITTSLSLATVPSTDSIQVOADARIFVINEMELEVSSGFNVDFAGBREVLI INSAGPIVDTEYGVALHPLOHGASYILI"

CDS

**ORIGIN**

Alignment Scores:  
Pred. No.: 9,5e-31 Length: 918  
Score: 420.50 Matches: 119  
Percent Similarity: 44.74% Conservative: 30  
Best Local Similarity: 35.74% Mismatches: 73  
Query Match: 24.01% Indels: 111  
DB: 8 Gaps: 14

US-09-787-737-2 (1-325) x BT005296 (1-918)

**Dy** 75 lleAnProArGlucduledlnArglleArgleugLeuglutyrglyGln 94  
:::|||||  
4 GTGAATCTCCAAGAGAGAGATCAGGAGGTAGGTCAGCTTCAAGATAACGCCAA 63

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

## AUTHORS

1 (bases 1 to 807)

Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

## TITLE

Arabidopsis ORF clones

## JOURNAL

## REFERENCE

## AUTHORS

2 (bases 1 to 807)

Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

## Direct Submission

## JOURNAL

## COMMENT

Submitted (17-APR-2003) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

## FEATURES

## source

Location/Qualifiers  
1..807  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="U60620"  
/ecotype="Columbia"  
/note="This clone is in PUNI 51"

## CDS

1..807  
/note="unknown protein"  
/codon\_start=1  
/product="At5g17810"  
/protein\_id="AAP12881.1"  
/db\_xref="GI:30017295"  
/translation="MNGEGASHSPSTSTPEVRAWSPKPEQIILLISFNSTVNP  
KDETVIRKMLEKFGVGDVNFYFQNRSSRRRRHQLAAATTAATSGADHGH  
MTAMSMHQYPCSNNEIDGFCNSLNFSSSQIPSPFGLFSSSSGGCENN  
GMENLFMYGHSDHNDQHHSSNAASVLNPSQNSQYEQEGFTVTFVGMVEV  
TKGALDKMTFGDDSVLLHSSGLPLPTDFGFLMHSQHGQTYFLVPRQT"

## ORIGIN

Alignment Scores: 2.45e-28 Length: 807  
Pred. No.: 394.50 Matches: 105  
Score: 47.93% Conservative: 38  
Percent Similarity: 35.12% Mismatches: 101  
Best Local Similarity:

Query Match: 22.53% Indels: 55  
DB: 8 Gaps: 9  
US-09-787-737-2 (1-325) x BT006232 (1-807)  
QY 41 SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60  
DB 19 TCACATAGCCCATCTCCACITCCACCGCAACAGTCGGGACAGTGTGGTCACTAACCG 78  
QY 61 GluGlnIleArgIleLeuSerIlePheAsnSerGlyThrIleAsnProArgGlu 80  
DB 79 GAGCAAAATCTTCATCTACTCGAATCCATCTTCAACAGTGGTACTGTAAACCCACAAAGAT 138  
QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaSerVal 100  
DB 139 GAAACGGTGCAGATGAAGAAGATGCTTGAGAAATTCGGTGTGGGAGACGCAACGTC 198  
QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120  
DB 199 TTCTACTGGTTTCAAAACCGACGTCAGATCTCGC-----CGGAGACACCGCGCAG 249  
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140  
DB 250 CTTTTCAGCAGCCACACCGCAGCCGCTCCATAGGAGCTGAGAACCCACCGACACATG 309  
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160  
DB 310 ACGGCCATGAGCATGATGATCAA-----TATCCTTGCAGCAACCAACGAGATT----- 354  
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerSerAlaGlnSerAla 180  
DB 355 -----GAT 357  
QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200  
DB 358 TTGGGGTTTGG-----AGTTGT 375  
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219  
DB 376 AGCACTTATCACTAATTAATCTTCTTAATGGATCGTCGTCATCTCAATCTCTCTTT 435  
QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyAsnCys 232  
DB 436 TTCCTCGGCTCTCTTCTTCAAGTGGTGGATGTGAGAACCAACATGTTATGGAGAATCTC 495  
QY 233 PheProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252  
DB 496 TTCAAAATGATGGCCATGAATCTGATCATATCATCAGCAGCAGCATCATAGTCAAT 555  
QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267  
DB 556 GCTGCATCAGTTTAAACCCATCTGATCAAACTCCAACTCCCAATACGAACAAGAGGG 615  
QY 268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286  
DB 616 TTTATGACCGGTGTTTATAAAACGGGAGTTCTCTATGGAAGTAAACAAAGAGCAATAGACATG 675  
QY 287 LysAlaAlaPheGlyAsnAspAlaValIleAsnSerPheGlyGlnProIleLeuThr 306  
DB 676 AAAACAATGTCGGTGATGATTCGGTGTACTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 735  
QY 307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325  
DB 736 GATGAGTTTGGTTTCTTCATGATCTCTTTTACACATGACGACAACTATTTCTCTGTA 792  
RESULT 10  
AK117910  
LOCUS  
DEFINITION  
Arabidopsis thaliana At5g17810 mRNA for unknown protein, complete  
cds, clone: RAFL19-09-D18.  
ACCESSION  
AK117910  
VERSION  
AK117910.1 GI:26450881  
KEYWORDS  
FLI\_CDNA; CAP trapper.

SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORS  
Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.  
Arabidopsis thaliana full-length cDNA  
Published Only in Database (2002)  
2 (bases 1 to 996)

TITLE  
JOURNAL  
Arabidopsis thaliana full-length cDNA  
Kawai, J., Hayashizaki, Y. and Shinozaki, K.  
Direct Submission  
Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: mseki@gsc.riken.go.jp).  
URL: <http://pfweb.gsc.riken.go.jp>, Tel: 81-45-503-9625, Fax: 81-45-503-9586

COMMENT  
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and Sali. This clone is in a modified pBluescript vector.  
Please visit our web site (<http://pfweb.gsc.riken.go.jp/>) for further details.

FEATURES  
Location/Qualifiers  
1..996  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/cultivar="Colombia"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="RAF119-09-D18"  
/note="common name: thale cress"  
1..996  
/gene="At5g17810/MVA3\_160"  
74..880  
/gene="At5g17810/MVA3\_160"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="BAC42548.1"  
/gb\_xref="GI:26450882"  
/translation="MNQGGASHSPSTPEVRAWSPKPEQILLESIFNSGTVNPP KDEVIRKMLEKFGAGVDANVYFQNRSSRRHRQLLAATTAATTAATSGAEDHQH MTAMSHQYPCSNNEIDIFGSCSNLSANYFLNGSSSQIFSLFLGSSGGCCNNH GMENLFKYGHSNHHQOHSSNAALNPDSQNSQVEQEGMTVFINGVPEV TKGALDMKTMFGDDSVLLHSSGLPLTDFEFGFMSLQHGQTYFLVPRQT"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.21e-28 Length: 996  
Score: 394.50 Matches: 105  
Percent Similarity: 47.83% Conservative: 38  
Best Local Similarity: 35.12% Mismatches: 101  
Query Match: 22.53% Indels: 55  
DB: 8 Gaps: 9

US-09-787-737-2 (1-325) x AK117910 (1-996)

Qy 41 SerSerPheSerAspArgileProAspProlYsProArgTtpAsnProLysPro 60  
Db 92 TCACATAGCCCATCTCCATCTCCACCGAACAGTCGGGACGCTGGTCACTAAACCG 151

Qy 61 GluGlnIleArgileLeuGlnSerIlePheAsnSerGlyThrIleAsnProArgGlu 80  
Db 152 GAGCAATCTTGACTCGAATCCATCTTCAACAGTGGTACTGTAAACCCCAAAAGAT 211

Qy 81 GluIleGlnArgileArgGluGlnGlyThrGlyGlnIleGlyAspAlaAsnVal 100  
|||

1212 GAACGGTGAGGATAAGAAAGATCTTGAGAAATTCGGTGTGTGGGAGACGCAACCTC 271  
Qy 101 PheTyrTtpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLys 120  
|||||  
Db 272 TTCTACTGGTTTCAAAACCGAGGTCAAGATCTCGC-----CGAGACACGGGAG 322

Qy 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140  
|||  
Db 323 CTTTTCAGCAGCCACCGCAGCGCCACCTCCATAGGAGCTGAAGACCCACGACACATG 382

Qy 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160  
|||  
Db 383 ACGCCATGAGCATGCGATCAA-----TATCCTTCGACCAACACGAGATT----- 427

Qy 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180  
|||  
Db 428 -----GAT 430

Qy 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200  
|||||  
Db 431 TTGGGGTTTGG-----AGTTGT 448

Qy 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219  
|||  
Db 449 AGCACTTATACGCTAATTAATTAATGATGCGTCTGATCTCAAAATCCCTCTCTTT 508

Qy 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyGlyAsnCys 232  
|||  
Db 509 TTCCTCGGCTCTCTTCTTCAAGTGGTGGATGAGAGAACAAATGGTATGGAGATCTC 568

Qy 233 PheProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252  
|||  
Db 569 TTCAAAATGATAGGCCATGAATCTGATCATAATCATCAGCAGCATCATAGTCAAAT 628

Qy 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267  
|||  
Db 629 GCTGCATCAGTTTAAACCCATCTGATCAAACTCCAACTCCCAATACCAAGAGGG 688

Qy 268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286  
|||  
Db 689 TTTATGACGGTGTATTAAACGGAGTTCTCTGGAAGTAAACAAAGAGCAATAGACATG 748

Qy 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306  
|||  
Db 749 AAAACAATGTCGGTGATGATTCCGGTGTTACTTCATCTCTGGTCTTCTTCCACT 808

Qy 307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325  
|||  
Db 809 GATGAGTTGGTTTCTTGATGATCTTTTACAACTAGCAAACTTATTTCTCTGTA 865

RESULT 11  
AC124143  
LOCUS 159878 bp DNA linear HTG 11-JUN-2002  
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 5 clone OSUNB0053D02, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 7 ordered pieces.  
AC124143  
AC124143.1 GI:21389220  
VERSION HTG; HTGS PHASE2  
KEYWORDS Oryza sativa (japonica cultivar-group)  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 159878)  
AUTHORS Chow, T.-Y., Hsing, Y.-I., Chen, C.-S., Chen, H.-H., Liu, S.-M., Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R., Chen, F.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H., Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C., Liu, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y., Yu, S.-W., Wu, H.-P. and Shaw, J.-F.  
TITLE Oryza sativa BAC OSUNB0053D02 genomic sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 159878)

```
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
```

*		1	12847:	contig of 12847 bp in length
*	12848		12947:	gap of unknown length
*	12948		67635:	contig of 54688 bp in length
*	67636		67735:	gap of unknown length
*	67736		84732:	contig of 17057 bp in length
*			84832:	gap of unknown length
*	84793		105305:	contig of 20413 bp in length
*	84893		105405:	gap of unknown length
*	105306		116399:	contig of 10994 bp in length
*	105406		116400:	gap of unknown length
*			116499:	gap of unknown length
*	116500		143095:	contig of 26596 bp in length
*	143096		143195:	gap of unknown length
*	143096		153589:	contig of 10394 bp in length
*	153590		153689:	gap of unknown length
*	153690		182868:	contig of 29179 bp in length.

				Location/Qualifiers
FEATURES				1..182868
source				

Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y., Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Harada, C., Ichishita, S., Honda, M., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikemura, K., Kamiya, K., Itoh, S., Itoh, T., Itoh, Y., Itoh, Y., Iwabuchi, A., Kaniwa, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Miyabayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saij, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshilara, R., Yukawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y., Yano, M., Jiang, J. and Gojobori, T.

The genome sequence and structure of rice chromosome 1  
Nature 420 (6913), 312-316 (2002)

22373376  
12447438

2 (bases 1 to 182681)

Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (13-JUN-2001) Takuji Sasaki, National Institute of  
Agrobiological Sciences, Rice Genome Research Program; Kamondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: tsasaki@nias.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>,  
Tel: 81-298-38-7441, Fax: 81-298-38-7468)

On Apr 16, 2002 this sequence version replaced gi:16904697.  
GENES were predicted from the integrated results of the following:  
GENSCAN1.0, ELASTN2.0, BLASTX2.0 as well as SplicePredictor  
(October 1998 version). The genomic sequence was searched against  
NCBI NonRedundant protein database, nr  
(<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA sequence database a  
RGP. Protein homologies of the coding regions were searched against  
NCBI NonRedundant Protein database using BLASTN 2.0 with the  
the identified cDNA sequences using BLASTN 2.0 with the  
corresponding DBJ accession no. and RGP clone ID.  
A gene with identity or significant homology to a protein is  
classified based on the protein name to indicate the homology level,  
such as same name, 'putative', and '-like protein'. A gene without  
significant homology to any protein but with EST homology (covering  
almost the entire length of partial sequence) is classified as an  
'unknown' protein. A gene predicted with a gene prediction program  
is classified as a 'hypothetical' protein.  
The orientation of the sequence is from -21M13 to M13rev of the  
BAC clone. This sequence of OSUNBB0063G05 clone has an overlap with  
OSUN003E08 (DBJ:AP003222) clone at the position 1 to 7853 of 5' end  
Detailed information on overlap and assembly quality together with  
narration of this entry is available at  
<http://rgp.dna.affrc.go.jp/GenomesSeq.html>.

Location/Qualifiers  
1. .182681  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="1"  
/clone="OSUNBB0063G05"  
complement(join(292..470,2341..2400,2526..2714,2941..3164..3273,3366..3453,3563..3796))  
/gene="OSUNBB0063G05.1"  
complement(join(292..470,2341..2400,2526..2714,2941..3164..3273,3366..3453,3563..3796))  
/gene="OSUNBB0063G05.1"  
/note="Contains ESTs AU032462(S10523), D46091(S10523)"  
/codon\_start=1  
/product="putative phosphatidic acid phosphatase beta"  
/protein\_id="BAB90492.1"  
/db\_xref="GI:20161571"  
LNA:EPHRRFGVAGMWTGPRPKMNTVPWAVTAVTGIVTFVYFRRRNVYDL  
HVAVLGILFSLVLTGLDTDAIKDAGRPVFWRCFPGDIAVFDNVTTGVCIGDAS  
VIXGKHKSFGSHTSWSFAGLWYSLAGKTVTFRRGHGKALCVLAPLVAAMVA  
ISRVDYVHHWHDVCTGGLVGLVSWYLIQFFPAPSDKVKVAPROGSGHTRTGT





Db 65913 GATTTCGGTGTACTTGCAGTGGCGAGAGAGGTTCCGGACCCGAAGCGCGTGG 65954  
 QY 57 AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsn 76  
 Db 65953 AACCCGGCGCGGAGCAGATCCGATCTCGAGCGGATCTTCAATCCGGGATCAAC 65794  
 QY 77 ProProArgGluGluIleGlnArgIleArgIleLeuGlnGluTyrGlyGlnIleGly 96  
 Db 65793 CCGCGCGCGAGAGATCCCGCGATCCGATCGAGCTGCGAGGATAGCGCGAGTCCGC 65734  
 QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLeuArg 116  
 Db 65733 GACGCCAACGCTTCTACTGTTCTCAGAACCGCAAGTCCCGCTCCAAGAACAGTCCGC 65674

RESULT 14  
 AK063503  
 LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:001-116-B11, full insert sequence.  
 DEFINITION AK063503  
 ACCESSION AK063503.1 GI:32973521  
 VERSION FLJ\_CDNA; oligo-capping.  
 KEYWORDS Oryza sativa (japonica cultivar-group)  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1  
 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Negata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtomo, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamura, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Sato, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
 Science 301 (5631), 376-379 (2003)  
 22752273  
 12869764

2 (bases 1 to 1282)  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 AUTHORS

# TITLE JOURNAL

Direct Submission  
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602 Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)  
 This clone is one of the 28K full-length cDNA clones from japonica rice.  
 URL : http://cdna01.dna.affrc.go.jp/cDNA/

## COMMENT

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Ohtomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamura, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Hara, A., Hashidume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Imotani, K., Ishii, Y., Kishikawa-Hirozane, T., Kojima, Y., Inamura, K., Imotani, K., Ishii, Y., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaka, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.

## FEATURES

### Location/Qualifiers

1..1282  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="001-116-B11"

### ORIGIN

#### Alignment Scores:

Pred. No.:	1..116-22	Length:	1282
Score:	338.00	Matches:	89
Percent Similarity:	46.34%	Conservative:	44
Best Local Similarity:	31.01%	Mismatches:	92
Query Match:	19.30%	Indels:	62
DB:	8	Gaps:	9

US-09-787-737-2 (1-325) x AK063503 (1-1282)

QY	53	LysProArgTrpAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSer	72
Db	171	AGGTCGGCGTGGACCGCGAAGCGGAGCAGATCTCATCTCTGGAGTCCATCTTCAACAGC	230
QY	73	GlyThrIleAsnProArgGluGluIleGlnArgIleArgLeuGlnGluTyr	92
Db	231	GGCATGTGTAAACCGCCCGAGGAGCGACCGTCCGCATCCGCAAGCTGCTCGAGCGTTC	290
QY	93	GlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLys	112
Db	291	GGCGCCCTCGCGAGCGCAACGCTTCTACTGTTTCCAGAACCCCGCTCGCTCCCGC	350
QY	113	HisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIle	132
Db	351	CGCCCGCAGCGC-----	362
QY	133	ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrPro	152
Db	363	---CAGTCGAGCGCGAGCGCGAGCGCGCGCGCGCTCGTCGCGATCTCTCTCCG	419

[illegible]

Query Match:	19..30%	Indels:	62
DB:	8	Gaps:	9
US-09-787-737-2 (1-325) x AK073232 (1-1378)			
QY	53	LysProArgTrpAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSer	72
Db	163	AGGTGCGGTGTGACGCGCGAAGCGGAGCAGATACTCATCTGGAGTCCATCTTCACAGC	222
QY	73	GlyThrIleAsnProProArgGluGluIleGlnArgIleArgIleArgLeuGlnGluTyr	92
Db	223	GGCATGGTGAACCCGCCCAAGCAGCAGACCGTCCGCATCCGAAGCTGCTCGAGCGCTTC	282
QY	93	GlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLys	112
Db	283	GGCGCGGTGCGGAGCGCCACACTCTTCTACTGTTCCAGAACCGCGCTCGCGCTCCCGC	342
QY	113	HisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIle	132
Db	343	CGCCGCCAGCGC-----	354
QY	133	ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrPro	152
Db	355	---CAGCTGCGAGCGCAGCGCGCAGCGCGCGCGCGCTCGTCTGGGATCTCTCCG	411
QY	153	ValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHis---	171
Db	412	-----ACTGCTGCTCGGTGCGCTCGCGCTCGCGCTCGCGCAGCC	447
QY	172	AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValValProValVal	191
Db	448	GGCTGCGCGGTTCGTCTCGGGATGTTCCGCGCAGCGC-----	486
QY	192	ThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGlu	211
Db	487	---GCCGCCGCGGTACAGCTCTCGTCTCTCATCTGCGCGCTCC-----	528
QY	212	AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAsp-----GlyAsnGly	228
Db	529	-----TCGCCGCGCTCGGTGGGATGATGATGGGGACGTGACTACGGGGCGGC	579
QY	229	GlyGlyAsnCysPheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys	248
Db	580	GGCGACGACCTGTTC-----GCCATCTCGAGCGAGATGGGGTAC	618
QY	249	ArgAspValGly-----LeuSerGlyGlyGlyAspValGlyAspAsn	262
Db	619	ATGACGCGCGCGCGCGCTCGTCTGCGCGCGCGCGGTGAGATCATCAGCAGCAGCAG	678
QY	263	ValTyr-----ProValArgMetThrValPheIleAsnGluMetProIleGlu	278
Db	679	CTCTACACCTGTCGTCACACCTGCGCATGACCGTGTTCATCAACGAGTGGCGAG	738
QY	279	ValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValLeuIleAsn	298
Db	739	GTSCCAAGGGGACCGATCGATCTGAGATCAATGTTTGGGCGAGGACGTGATGCTGCTGCAT	798
QY	299	SerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsn	318
Db	799	TCAACGGGTGCTCTTCTCCAGCCACGAGTACGGCATCTCTCTCCATCTCTCCAGATG	858
QY	319	GlyAlaIleTyrTyrLeuIle	325
Db	859	GGCGAGAGCTACTTCTCTGTC	879

Search completed: September 10, 2004, 10:47:34  
Job time : 4872 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 10, 2004, 06:25:41 ; Search time 471 Seconds  
(without alignments)  
2931.347 Million cell updates/sec

Title: US-09-787-737-2

Perfect score: 1751

Sequence: 1 MSSSNKNWPSMFKSKPCNNN.....TDFGVGYQLQNGAIYYLI 325

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Rgapop 6.0 , Rgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cgn2.1/USPTO spool h/US0978737/runat 01092004 112454 23588/app query.fasta 1.519  
-PB=N Geneseq 29Jan04 -QWNT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-NODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US0978737 @CGN 1 1 470 @runat 01092004 112454 23588 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002s:\*
- 7: Geneseqn2003as:\*
- 8: Geneseqn2003bs:\*
- 9: Geneseqn2003cs:\*
- 10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1751	100.0	1214	5 AAF59786	Arabidops
2	645.5	36.9	1518	5 AAF59787	Arabidops
3	623	35.6	1502	9 ADD30297	Plant yie
4	394.5	22.5	807	6 ABK65270	Arabidops
5	394.5	22.5	807	9 ADD30133	Plant yie
6	342	19.5	994	6 ABK65367	Arabidops
7	328.5	18.8	1367	4 AAS02222	Corn Wusc
8	212.5	12.1	816	9 ADD30299	Plant yie

9	212.5	12.1	1326	3 AAC38580	Arabidops
10	200	11.4	1209	3 AAC47056	Arabidops
11	189.5	11.4	1227	9 ADD30291	Plant yie
12	188	10.7	371	3 AAC56282	Pinus rad
13	187.5	10.7	1338	4 AAS02219	Corn Wusc
14	187	10.7	741	4 AAS02227	Soybean W
15	175.5	10.0	844	4 AAS02226	Soybean W
16	173	9.9	904	4 AAS02229	Soybean W
17	172.5	9.9	906	4 AAS02228	Soybean W
18	171.5	9.8	522	7 AAD47508	Arabidops
19	171.5	9.8	648	3 AAS5917	Eucalyptu
20	171.5	9.8	717	6 ABQ65592	Arabidops
21	170	9.7	506	4 AAS02225	Soybean W
22	170	9.7	1190	2 AAZ20289	Aspen hom
23	167.5	9.6	720	4 AAS02220	Corn Wusc
24	167.5	9.6	783	6 ABK65267	Arabidops
25	167.5	9.6	783	9 ADD30617	Plant yie
26	167.5	9.6	783	9 ADE31512	Plant yie
27	167.5	9.6	1136	2 AAZ20288	Aspen hom
28	167	9.5	945	3 AAC42316	Arabidops
29	166	9.5	767	4 AAS02221	Corn Wusc
30	140.5	8.0	1194	4 AAS02224	Corn Wusc
31	135.5	7.7	513	4 AAS02223	Corn Wusc
32	129	7.4	441	3 AAG55958	Eucalyptu
33	117	6.7	4791	8 ADA30129	DNA encod
34	113	6.5	771	6 ABS63002	Selected
35	113	6.5	771	6 ABT11607	Yeast sel
36	112.5	6.4	1302	4 AAD06478	Arabidops
37	112.5	6.4	1302	9 ADR37278	Plant yie
38	111.5	6.4	7377	7 ACF69433	Phototrab
39	111.5	6.4	110000	7 ACF67367_23	Continuation (24 o
40	111.5	6.4	110000	7 ACF65386_5	Continuation (6 of
41	111	6.3	2287	6 ABA95489	Human ker
42	109.5	6.3	462	4 AAS02218	Corn Wusc
43	108.5	6.2	3082	9 ADB62521	Human cDN
44	108.5	6.2	3845	7 ACC72452	Human sec
45	108.5	6.2	3886	6 ABQ75807	Human MDD

## ALIGNMENTS

RESULT 1

AAF59786

ID AAF59786 standard; cDNA; 1214 BP.

XX

AC AAF59786;

XX

DT 04-MAY-2001 (first entry)

XX

DE Arabidopsis thaliana M6 homeodomain protein cDNA.

XX

KW Homeodomain-like; homeobox gene; M6; plant differentiation; MSH;  
KW many shoot; expression control; adventitious bud formation;  
KW branching induction; cytokinin independent; transgenic plant; crop;  
KW agriculture; ss.

XX

OS Arabidopsis thaliana.

XX

PN WO200107618-A1.

XX

PD 01-FEB-2001.

XX

PF 21-JUL-2000; 2000WO-JP004904.

XX

PR 22-JUL-1999; 99JP-00207995.

XX

PA (SUNR ) SUNTORY LTD.

XX

PI Kakimoto T;

XX

DR WPI; 2001-182796/18.

XX

DR P-PSDB; AAB60642.

XX

PT Arabidopsis thaliana originated homeobox genes encoding proteins  
 PT participating in differentiation with ability to promote adventitious bud  
 PT formation and branching induction, applicable in improving plant cells  
 PT and plants e.g. for crops.

XX  
 XX  
 PS Claim 3; Page 20-23; 34pp; Japanese.

XX  
 CC The invention relates to two novel Arabidopsis thaliana proteins M6 and  
 CC M8 (AA60642 and AA60643, respectively), and to the cDNAs encoding them  
 CC (AA60642 and AA60643). The proteins contain homeodomain- like sequences  
 CC and participate in differentiation, having the ability to promote  
 CC adventitious bud formation and branching induction by controlling  
 CC expression of the many shoot (MSH) gene without the requirement for  
 CC cytokinin. The invention also relates to expression vectors and host  
 CC cells which contain M6 or M8 nucleic acid sequences, the recombinant  
 CC production of the M6 or M8 protein, a plant or plant cell transfected  
 CC with the M6 or M8 gene, and methods for inducing differentiation,  
 CC adventitious bud formation and/or plant branching via expression of the  
 CC transacted M6 or M8 gene. Nucleic acids encoding the homeodomain  
 CC proteins may be used to generate transgenic plants with improved  
 CC adventitious bud formation and branching for use e.g., as agricultural  
 CC crops. The present sequence represents cDNA encoding the Arabidopsis  
 CC thaliana M6 protein

XX  
 SQ Sequence 1214 BP; 395 A; 228 C; 226 G; 365 T; 0 U; 0 Other;

Alignment Scores: 1.13e-180 Length: 1214  
 Pred. No.: 1751.00 Matches: 325  
 Score: 1751.00 Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 5

US-09-787-737-2 (1-325) x AAF59786 (1-1214)

QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsn 20  
 DB\* 36 ATCTCTCTCTCAAAACAAATTTGCCAAGCATGTTCAAAATCCAAACCTTGCAACAATAAT 95  
 QY 21 HisHisGlnHisGlnLysAsnThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40  
 DB 96 CATCATCATCAACATGAATTCGATACCTCCATCTTACATGCACTACTCTAATTGCAACCTA 155  
 QY 41 SerSerSerPheSerSerAspArgLysProAspProLysProArgTrpAsnProLysPro 60  
 DB 156 TCAATCTCTCTCTCTCAGATCGATACCGATCCATCTTAAACCGAGATGGAATCTCTAAACCG 215  
 QY 61 GluGlnLysArgLysLeuGluSerLysPheAsnSerGlyThrLysAsnProProArgGlu 80  
 DB 216 GAGCAGATTAGGATACCTGAATCAATCTTCAATTCGGGTACTATTAACCCACCTAGAGAG 275  
 QY 81 GluLysGlnArgLysArgLysGlnGluLysGlnLysGlnLysGlnLysGlnLysGlnLys 100  
 DB 276 GAGATTCAAGAAATCCCGATCCCGCTTCAAGAAATGTCATAATCGGTGACGCAACCGTG 335  
 QY 101 PheTyrTrpPheGlnAsnArgLysSerArgLysHisLysLysLeuArgValHisHisLys 120  
 DB 336 TTTTACTGTTTCAAAACCGGAATCTCGACAAACCAAGCTTCGTGTTTCATCACAAA 395  
 QY 121 SerProLysMetSerLysLysAspLysThrValLysProSerThrAspAlaAspHisCys 140  
 DB 396 AGCCCTAAATGTCAAGAAAGACAGACGCTTATCTCTAGTACTGACGCTGATCATTTGT 455  
 QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160  
 DB 456 TTGTGTTTGTAAACCAAGAACCGGATATATCCGGTTCAAAACAATGAGTTGTGTGTA 515  
 QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180  
 DB 516 ACCGAAACCGCGCGTTTCTTATTTCCGGTTTCATATGATGATCCGAGCGCTCTCAATCAGCG 575  
 QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200

DB 576 TTGGTTTTCGCGATTTTGTGTACCGGTGTAACGGAAGAGGATGGCATTTCTTACC 635  
 QY 201 ValAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsn 220  
 DB 636 GTTAATAACGCGTTAATTTGGAGACTAACGAAAAATTTGTATAAAATTCGCGCATCAAT 695  
 QY 221 LeuTyrGlyAspGlyAsnGlyGlyGlyAsnCysPheProProLeuThrValProLeu 240  
 DB 696 TTATACGCGGAGATGGAAATGGCGGTGGAATTTGTTTCTCTTCTGACTGTTCATTA 755  
 QY 241 ThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyGluAspValGly 260  
 DB 756 ACCATCAATCAATCTCAAGAAAAACGAGATGTAGGATTTATCCGGTGGTGAACGCTCGGA 815  
 QY 261 AspAsnValTyrProValArgMetThrValPheIleAsnGluMetProIleGluValVal 280  
 DB 816 GATATATGTTTATCCGGTGAGATGACGGTGTATTATTAACGAGATGCTTATCGAAGTAGTG 875  
 QY 281 SerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPhe 300  
 DB 876 TCTGATTTATTTCAACGTTAAGGACGCTTTCGAAAACGATGCGGTTTGTATCAACTCGTTT 935  
 QY 301 GlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAla 320  
 DB 936 GGCCAGCGCTATTTCTTACAGATGATTTGGTGTATTCTATCAACCTCTCCAAATGGCGCA 995  
 QY 321 IleTyrTyrLeuLeu 325  
 DB 996 ATCTATTATCTTATT 1010  
 RESULT 2  
 AAF59787  
 ID AAF59787 standard; cDNA; 1518 BP.  
 AC AAF59787;  
 XX  
 XX 04-MAY-2001 (first entry)  
 DT  
 DT Arabidopsis thaliana M8 homeodomain protein cDNA.  
 DE  
 XX Homeodomain-like; homeobox gene; M8; plant differentiation; MSH;  
 KW many shoot; expression control; adventitious bud formation;  
 KW branching induction; cytokinin independent; transgenic plant; crop;  
 KW agriculture; ss.  
 KW  
 XX Arabidopsis thaliana.  
 OS  
 XX  
 XX WO200107618-A1.  
 FN  
 XX  
 XX 01-FEB-2001.  
 PD  
 XX 21-JUL-2000; 2000WO-JP004904.  
 PF  
 XX 22-JUL-1999; 99JP-00207995.  
 PR  
 XX (SUNR ) SUNTORY LTD.  
 PA  
 XX Kakimoto T;  
 PI  
 XX WPI; 2001-182796/18.  
 DR P-PSDB; AAB60643.  
 DX  
 XX Arabidopsis thaliana originated homeobox genes encoding proteins  
 PT participating in differentiation with ability to promote adventitious bud  
 PT formation and branching induction, applicable in improving plant cells  
 PT and plants e.g. for crops.  
 XX  
 PS Claim 6; Page 25-28; 34pp; Japanese.  
 CC The invention relates to two novel Arabidopsis thaliana proteins M6 and  
 CC M8 (AA60642 and AA60643, respectively), and to the cDNAs encoding them  
 CC (AA60642 and AA60643). The proteins contain homeodomain- like sequences



and participate in differentiation, having the ability to promote adventitious bud formation and branching induction by controlling expression of the many shoot (MSH) gene without the requirement for cytokinin. The invention also relates to expression vectors and host cells which contain M6 or M8 nucleic acid sequences, the recombinant production of the M6 or M8 protein, a plant or plant cell transfected with the M6 or M8 gene, and methods for inducing differentiation, adventitious bud formation and/or plant branching via expression of the transfected M6 or M8 gene. Nucleic acids encoding the homeodomain proteins may be used to generate transgenic plants with improved adventitious bud formation and branching for use e.g., as agricultural crops. The present sequence represents cDNA encoding the Arabidopsis thaliana M8 protein

SQ Sequence 1518 BP; 425 A; 389 C; 279 G; 425 T; 0 U; 0 Other;

Alignment Scores:		
Pred. No.:	4.56e-60	1518
Score:	645.50	Matches: 156
Percent Similarity:	50.61%	Conservative: 42
Best Local Similarity:	40.39%	Mismatches: 84
Query Match:	36.96%	Indels: 119
DB:	5	Gaps: 17

US-09-787-737-2 (1-325) x AAF59787 (1-1518)

[illegible]

Qy	201	ValAsnAsnGly---	ValAsnLeuGluThrAsnGlu-	211
		:::       :::	:::	
Db	830	ATCATGAACGGTAGTGTGAGTTATGCAACTCATCAACAAC	CTTGAGTGAGAAAGAA	889
Qy	211	----	----	211
Db	890	GTTGAAGAAATGAGGATGAAGATGTTGCAACAGGCCACAGACTCAGATTGTTACGCTACC	949	
Qy	212	AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGlyGlyAsn	231	
Db	950	ACTAATCATCAATAAGCTTCTTACAAC	AAACAACAACAACAACAATAAC	997
Qy	232	Cys-----PheProPro-	----	235
Db	998	ATCATGCTTCATATTCTCCCACTACTTCTACTGCCACACTATTACTACTTCGCAATTCT	1057	
Qy	236	---LeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer	254	
Db	1058	CTCGTACTGTCCCATCAACTTCGGACCGAGTTCAGTTCAAGCGGAC-	1105	
Qy	255	GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu	274	
Db	1106	-----GCACGAATAAGAGTTTTCATCAATGAA	1132	
Qy	275	MetProIleGluValValSerGlyLeuPheAsnValLysAlaIlePheGlyAsnAspAla	294	
Db	1133	ATGGAGCTTGAAGTGAGCTCAGGCCGTTCAATGTGAGGATGCATTGGGGAAGAGTTT	1192	
Qy	295	ValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGln	314	
Db	1193	GTTCTGATTAAATCCCGGGGTACGCCCATGTACCCGATGAATATGCGCTGCTCTTCCAC	1252	
Qy	315	ProLeuGlnAsnGlyAlaIleTyrTyrLeuIle	325	
Db	1253	CCCTTTCAACACGAGCGCTCGTACTATCTCATC	1285	

### RESULT 3

ADD30297  
ID ADD3

XX  
AC

XX  
116706000

DT 15-JAN-2004 (first entry)

XX  
DE  
vv

ds; transcription factor; transgenic plant; growth  
seed germination rate; plant vigor; seedling vigor

XX OS *Arabidopsis thaliana*.

XX  
DN

PN WO2003013227-A2.  
XX

PD 20-FEB-2003.

09-AUG-2002: 2002WO-US025805

XX  
00 1100 200Z, 200ZMO-05073805.

PR 09-AUG-2001; 2001US-0310847P.  
PR 19-NOV-2001; 2001US-0336049P

11-DEC-2001; 2001US-0338692P;

PR 14-JUN-2002; 2002US-00171468.  
XX

PA (MEND-) MENDEL BIOTECHNOLOGY INC.

XX  
PI Ratcliffe O  
Piechmann II

Kacchilli O, Kiechmann JL, Adam LJ, Dubell AT, Heard JB;  
 Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, van C.

PI Broun PE;  
XX

DR WPI; 2003-248221/

DR P-PSDB; ADD30298.  
XX  
PT New plant transcription factor polynucleotides and polypeptides, useful

PT in producing transgenic plants with commercially valuable properties,  
 PT such as an alteration in a plant growth characteristic, e.g. growth rate  
 PT or apomixis.

XX Disclosure; SEQ ID NO 326; 454pp; English.

XX The invention relates to a number of isolated Arabidopsis thaliana cDNA  
 CC sequences and their encoded proteins which are especially transcription  
 CC factor related cDNA's and proteins. The isolated or recombinant plant  
 CC transcription factor polynucleotides and polypeptides are useful in  
 CC producing transgenic plants with commercially valuable properties, i.e.  
 CC modified or altered desirable traits as compared to a reference plant,  
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,  
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and  
 CC flower senescence. Sequence information related to the polynucleotides  
 CC and polypeptides can also be used in bioinformatic search methods. The  
 CC transgenic plant is useful for growing a progeny plant from a parent  
 CC plant. This sequence represents one of the cDNAs of the invention.

XX Sequence 1502 BP; 414 A; 390 C; 273 G; 425 T; 0 U; 0 Other;

Alignment Scores: Length: 1502  
 Pred. No.: 1,27e-57 Matches: 164  
 Score: 623.00 Conservative: 40  
 Percent Similarity: 49.64% Mismatches: 83  
 Best Local Similarity: 39.90% Indels: 124  
 Query Match: 35.58% Gaps: 18  
 DB:

US-09-787-737-2 (1-325) x ADD30297 (1-1502)

QY 1 MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsnAsn 20  
 DB 160 ATGGCTTCTCGATAGACACTGGCCAGCATGTTCAAGTCCAAACCT-----CATCC 213  
 QY 21 HisHisGlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeu 40  
 DB 214 CATCAATGGCAACATGATCACTCTCT-----CTCTGCTTCTGCTTCTCACC 267  
 QY 41 SerSerSerPheSerSer-----AspArgileProAspProLysProArgTirp 56  
 DB 268 TCTTCT 327  
 QY 57 AsnProLysProGluGlnLeuArgileGluLeuSerLysPheAsnSerGlyThrIleAsn 76  
 DB 328 AATCAAAAGCCAGAGAGATTCGATACCTTCAAGCAATCTTAACTCCGGATGGTGAAT 387  
 QY 77 ProProArgGluGluLeuGlnArgileArgileGluLeuGluTyrGlyGlnIleGly 96  
 DB 388 CCTCCAGAGAGAG-----ATCAGGCTTCAAGATACGGCCAGTCCGT 432  
 QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116  
 DB 433 GATGCTAACGCTTCTACTGTTTCAAAACCGTAAGTCCCGTAGTAAACACAAATCCGCG 492  
 QY 117 ValHisHis----- 119  
 DB 493 CTCCTCCCAACCACTCCAAACACTCTCTCTCTCAACCGCAACCGCAGCGCGCAA 552  
 QY 120 -----LysSerProLysMetSerLysLysAspLysThrValIlePro--- 133  
 DB 553 CTTTCGGCTTCT 612  
 QY 134 -----SerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGly 149  
 DB 613 AAAACCAAGAACAGAACACTAATCTCTCTTTGGGT---GGTAGTCAAAATGATGGGG 669  
 QY 150 LeuTyrProValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPhePro 169  
 DB 670 ARGTTTCCA-----CCGGAACCGCGG---TTTCTCTCTCCCG 702  
 QY 170 ValHisAsn-----AspProSerAlaAlaGlnSerAlaPheGlyPhe----- 183

DB 703 GTCTCCACTGTCGAGGGTTGAAGGATATACCGTCTCTCATCCCAATTAGGTTTCTCTCC 762  
 QY 184 GlyAspPheVal-----ValProValValThrGluGluGlyMetAlaPheSerThr 200  
 DB 763 GGTGATATGATTGAGCACAACAAACCGCTCCAAAGTGTACCGGACTCTCTGAGTGAG 822  
 QY 201 ValAsnAsnGly-----ValAsnLeuGluThrAsnGlu----- 211  
 DB 823 ATCATGAACGGTAGTGTGAGTTATGGAATCATCATCAACACACTTGAAGTGAAGA 882  
 QY 211 ----- 211  
 DB 883 GTTGAAGAAATGAGATGAAGATGTTGCAACAGCCACAGACTCAGATTTGTTACGCTACC 942  
 QY 212 AsnPheAspLysIleProAlaIleAsnLeuTyrGlyAspGlyAsnGlyGlyGlyAsn 231  
 DB 943 ACTAATCATCAATAGCTTCTTACAC-----AACACACACAAACAAATAAC 990  
 QY 232 Cys-----PheProPro----- 235  
 DB 991 ATCATGCTTCATATTCCTCTCCCACTTCTTACTGCCACCACTATTACTTCTCGATTCT 1050  
 QY 236 -----LeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer 254  
 DB 1051 CTCGCTACTGTCCTCCATCAACTTCGGACACGCTTCAAGTTCACAGCGAC----- 1098  
 QY 255 GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu 274  
 DB 1099 -----GCACGAATAAGAGTTTTTCATCAATGAA 1125  
 QY 275 MetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAla 294  
 DB 1126 ATGAGCTTGAAGTGAAGTCAAGCCGTTCAATGTGAGGATGCAATTTGGGAGAGGTT 1185  
 QY 295 ValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGln 314  
 DB 1186 GTTCTGATTAATTCGCGGTTCAGCCCATGTCCACGAATGAATATGGCGTCTCTTCAC 1245  
 QY 315 ProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325  
 DB 1246 CCTCTTCAACACGAGCGCTGATCTATCTGATC 1278  
 RESULT 4  
 ABK65270 standard; cDNA; 807 BP.  
 XX ABK65270;  
 XX 02-JUL-2002 (first entry)  
 XX Arabidopsis cDNA encoding a transcription factor #122.  
 XX Plant; ss; gene; transcription factor; transgenic; agriculture;  
 XX metabolic chemical; environmental stress; drought;  
 XX microbial disease resistance; herbicide resistance; seed yield;  
 XX fruit yield; growth rate; leaf senescence; flower senescence.  
 XX Arabidopsis thaliana.  
 XX WO200215675-A1.  
 XX 28-FEB-2002.  
 XX 22-AUG-2001; 2001WO-US026189.  
 XX 22-AUG-2000; 2000US-0227439P.  
 XX 16-NOV-2000; 2000US-00713994.  
 XX 18-APR-2001; 2001US-00837944.  
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 XX (PILG/) PILGRIM M.  
 XX (CREE/) CREELMAN R.  
 XX (DUBE/) DUBELL A J.

PA (HEAR/) HEARD J.  
 PA (JIAN/) JIANG C.  
 PA (KEDD/) KEDDIE J.  
 PA (ADAM/) ADAM L.  
 PA (RATC/) RATCLIFF O.  
 PA (REUB/) REUBER J L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (YUGG/) YU G.  
 PA (PINE/) PINEDA O.  
 XX  
 PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddle J;  
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
 XX  
 DR WPI: 2002-292022/33.  
 DR P-PSDB: AAU93084.  
 XX  
 An isolated or recombinant polynucleotide used to produce a transgenic plant.  
 XX  
 Claim 4; Page 512-513; 941pp; English.  
 XX  
 The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ectopic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemicals, pest tolerance, environmental stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf and flower senescence and many other traits listed in the specification). The present sequence is one of the 232 polynucleotides encoding an A. thaliana transcription factor  
 XX  
 SQ Sequence 807 BP; 234 A; 197 C; 168 G; 208 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 4,16e-33 Length: 807  
 Score: 394.50 Matches: 105  
 Percent Similarity: 47.82% Conservative: 38  
 Best Local Similarity: 35.12% Mismatches: 101  
 Query Match: 22.53% Indels: 55  
 DB: 6 Gaps: 9  
 US-09-787-737-2 (1-325) x ABK65270 (1-807)  
 QY 41 SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60  
 Db 19 TCACATAGCCCATCTCCACTTCCACCAACAGCGCGGACGGTGTGTCTACCTAAACCG 78  
 QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGlu 80  
 Db 79 GAGCAATCTTGATCTCGAATCCATCTTCAACAGTGTACTGTAAACCCACCAAAAGAT 138  
 QY 81 GluLeuGlnArgIleArgIleArgLeuGlnGlyGlnIleGlyAspAlaAsnVal 100  
 Db 139 GAAACGGTGAGATAAGAAAGATCTTGAGAAATTCGGTGTGTGGAGAGCAAAAGCTC 198  
 QY 101 PheYrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120

Db 199 TTCTACTGGTTTCAAAACCGCGGTCAAGATCTCGC-----CGGAGACACCGGAG 249  
 QY 121 SerProLysMetSerLysAspLysThrValIleProSerThrAspAlaAspHisCys 140  
 Db 250 CTTTATGAGCAGCCACCGCGCCACCTCCATAGAGCTGAAGACCCACGACACATG 309  
 QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValVal 160  
 Db 310 ACGGCCATGAGCATGCATCAA-----TATCCTTCAGCAACACGACGATTT----- 354  
 QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180  
 Db 355 -----CAT 357  
 QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200  
 Db 358 TTGGGGTTTGA-----AGTTGT 375  
 QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219  
 Db 376 AGCAACTTATCAGCTAATTAATTCCTTAATGGATCGTCGTCATCTCAATCCCTTCCTTT 435  
 QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyGlyAsnCys 232  
 Db 436 TTCCTCGGCTCTCTTCTTCAAGTGGTGTGTGAGACAAACAATGGTATGGAGAAATCTC 495  
 QY 233 PheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252  
 Db 496 TTCAAATATGTATGGCCATGCAATCTGATATAATCATCAGCAGCAGCATCATGCTCAAT 555  
 QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267  
 Db 556 GCTGCATCAGTTTAAACCCATCTGATCAAACTCCAACTCCCAATACGAACAGAGGG 615  
 QY 268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286  
 Db 616 TTTATGACGGTGTATTAACCGAGTTCCTGTGAAGTAAACAAAGGAGCAATGACATG 675  
 QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306  
 Db 676 AAAACAATGTCGGTGATGATTCGGTGTACTTCATCTCTGTCTTCTCTCTCTCCACT 735  
 QY 307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325  
 Db 736 GATGAGTTGGTTTCTTGATGCAATCTTTACACATGACAACTTATTTCTCTGTTA 792  
 RESULT 5  
 ADD30133  
 ID ADD30133 standard; cDNA; 807 BP.  
 XX  
 AC ADD30133;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Plant yield-related polynucleotide clone G1586.  
 XX  
 KW ds; transcription factor; transgenic plant; growth rate; senescence;  
 KW seed germination rate; plant vigor; seedling vigor.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FN WO2003013227-A2.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 09-AUG-2002; 2002WO-US025805.  
 XX  
 PR 09-AUG-2001; 2001US-0310847P.  
 PR 19-NOV-2001; 2001US-0336049P.  
 PR 11-DEC-2001; 2001US-0338692P.  
 PR 14-JUN-2002; 2002US-00171468.  
 XX

PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;  
 PI Pilgrim M, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;  
 PI Broun PE;  
 XX WPI; 2003-248221/24.  
 DR P-PSDB; ADD30134.  
 XX New plant transcription factor polynucleotides and polypeptides, useful  
 XX in producing transgenic plants with commercially valuable properties,  
 XX such as an alteration in a plant growth characteristic, e.g. growth rate  
 XX or apomixis.  
 XX Disclosure; SEQ ID NO 162; 454pp; English.  
 XX The invention relates to a number of isolated Arabidopsis thaliana cDNA  
 XX sequences and their encoded proteins which are especially transcription  
 XX factor related cDNA's and proteins. The isolated or recombinant plant  
 XX transcription factor polynucleotides and polypeptides are useful in  
 XX producing transgenic plants with commercially valuable properties, i.e.  
 XX modified or altered desirable traits as compared to a reference plant,  
 XX such as an alteration in a plant growth characteristic, e.g. growth rate,  
 XX germination rate of seeds, vigor of plants and seedlings, or leaf and  
 XX flower senescence. Sequence information related to the polynucleotides  
 XX and polypeptides can also be used in bioinformatic search methods. The  
 XX transgenic plant is useful for growing a progeny plant from a parent  
 XX plant. This sequence represents one of the cDNAs of the invention.  
 XX SQ Sequence 807 BP; 234 A; 197 C; 168 G; 208 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4,16e-33 Length: 807  
 Score: 394.50 Matches: 105  
 Percent Similarity: 47.83% Conservative: 38  
 Best Local Similarity: 35.12% Mismatches: 101  
 Query Match: 22.53% Indels: 55  
 DB: 9 Gaps: 9

US-09-787-737-2 (1-325) x ADD30133 (1-807)

QY 41 SerSerSerPheSerAspArgIleProAspProLysProArgTyrAsnProLysPro 60  
 DB 19 TCACATAGCCATCTCTCCATTCACCGAACCCAGTCCGGGACCGTGGTCACTTAACCG 78  
 QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProArgGlu 80  
 DB 79 GAGCAATCTTGATACGATCGATCCATCTTCAACAGTGTGTACTGTTAACCCCAAAAGAT 138  
 QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100  
 DB 139 GAAACGGTGGAGATAAGAAAGATGCTTCAGAAATTCGGTGTGTGGGAGACGCAACGTC 198  
 QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLys 120  
 DB 199 TTCTACTGTTTCAAAACCGGACGTCAGATCTGC-----CGGAGACACCGCAG 249  
 QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140  
 DB 250 CTTTATAGAGCCACCCAGCCGCCACCTCCATAGGAGTGAAGACACCCAGCACATG 309  
 QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160  
 DB 310 ACGGCCATGAGCATGCATCAA-----TATCCTTGCAGCAACACGAGATT----- 354  
 QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180  
 DB 355 -----GAT 357  
 QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200  
 DB 358 TTGGGGTTTGA-----AGTTGT 375

QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219  
 DB 376 AGCAACTATCAGCTAATTAATCTTCTTAATGATCGTCTCATCTCAATCCCTTCCTTT 435  
 QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyGlyAsnCys 232  
 DB 436 TTCTCTCGGCTCTCTCTTCAAGTGGTGGTGTGAGAACCAACATGGTATGGAGATCTC 495  
 QY 233 PheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252  
 DB 496 TTCAAATGTATGCCCATCATCTCATATAATCATCAGCAGCAGCATCATAGTCAAAAT 555  
 QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267  
 DB 556 GCTGCATCAGTTTAAACCCATCTCATCAAACTCAATCCCAATACGAAACAAGGG 615  
 QY 268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286  
 DB 616 TTTATGACGGTGTATATAAACCGAGTTCCTATGGAAGTAAACAAAGGAGCAATAGACATG 675  
 QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306  
 DB 676 AAAACAATGTCGGTGAATTCGGTGTACTTCTCTCTGCTTCTCTCTCTCTCCACT 735  
 QY 307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325  
 DB 736 GATGAGTTTGGTTCTTGATGATCTTTTACAACTGAGCAAACTTATTTCTCTGGTA 792  
 RESULT 6  
 ABK65367  
 ID ABK65367 standard; cDNA; 994 BP.  
 XX AC ABK65367;  
 XX XX 02-JUL-2002 (first entry)  
 XX DE Arabidopsis cDNA encoding a transcription factor #219.  
 XX KW Plant; ss; gene; transcription factor; transgenic; agriculture;  
 KW metabolic chemical; environmental stress; drought;  
 KW microbial disease resistance; herbicide resistance; seed yield;  
 KW fruit yield; growth rate; leaf senescence; flower senescence.  
 XX Arabidopsis thaliana.  
 XX OS W0200215675-A1.  
 XX PN 28-FEB-2002.  
 XX PD 22-AUG-2001; 2001WO-US026189.  
 XX PF 22-AUG-2000; 2000US-0227439P.  
 XX PR 16-NOV-2000; 2000US-00713994.  
 XX PR 18-APR-2001; 2001US-00837944.  
 XX XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (PILG/) PILGRIM M.  
 PA (CREE/) CREELMAN R.  
 PA (DUBE/) DUBELL A J.  
 PA (HEAR/) HEARD J.  
 PA (JIAN/) JIANG C.  
 PA (KEDD/) KEDDIE J.  
 PA (ADAM/) ADAM L.  
 PA (RATC/) RATCLIFF O.  
 PA (REUB/) REUBER J L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (YUGG/) YU G.  
 PA (PINE/) PINEDA O.  
 XX Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;  
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
 XX WPI; 2002-292022/33.

DR

P-PSDB; AAU93181.

XX An isolated or recombinant polynucleotide used to produce a transgenic plant.

PT Claim 4; Page 889-891; 941pp; English.

XX The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ectopic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemicals, pest tolerance, environmental stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf and flower senescence and many other traits listed in the specification). The present sequence is one of the 232 polynucleotides encoding an A. thaliana transcription factor

SQ Sequence 994 BP; 281 A; 209 C; 205 G; 299 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,95e-27 Length: 994  
Score: 342.00 Matches: 87  
Percent Similarity: 44.22% Conservative: 47  
Best Local Similarity: 28.71% Mismatches: 95  
Query Match: 19.53% Indels: 75  
DB: 6 Gaps: 4

US-09-787-737-2 (1-325) x ABK65367 (1-994)

QY 23 HisGlnHisGluIleAspThrProSerTyrMetHisTyrSerAsnCysAsnLeuSerSer 42  
DB 11 AACAAACACCATAGCCCAACCCCATAGTCGCTCACCCCATCC-TCGCGCTCCGGT 69  
QY 43 SerPheSerAspArgIleProAspProLysProArgTyrAsnProLysProGluGln 62  
DB 70 TCCACCTCAGCAGAACCGGT-----CGGTCCGATGTCACCTTAACCGAACA 120  
QY 63 IleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluIle 82  
DB 121 ATACTCATCTAGTGCATCTCCACAGTGTATGTTAACTCCCAAGAGAGACG 180  
QY 83 GlnArgIleArgIleArgLeuGlnGluTyrGlnIleGlyAspAlaAsnValPheTyr 102  
DB 181 GTAAGGATAGAAGATGCTCGAATAATTTCGCCGCGTGGGAGATCAAAATGCTCTAT 240  
QY 103 TrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerPro 122  
DB 241 TGGTTTCAAAACCGCGGTCAAGTCCCGTCGGAGACGCA----- 282  
QY 123 LysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGly 142  
DB 282 ----- 282  
QY 143 PheValAsnGlnGluThrGlyLeuTyrProValClnAsnAsnGluLeuValThrGlu 162  
DB 282 ----- 282

QY 163 ProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGly 182  
DB 283 -----CAGCTACAGGCTGCAGCTCA 303  
QY 183 PheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThrValAsn 202  
DB 304 GCAGCGGAC-----GCAACCAACCAACACTTCTAGCTCTTCTTCTTCTATGGT 351  
QY 203 AsnGlyValAsnLeuLeuThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyr 222  
DB 352 GGTGGATGCTATATCAAGCAATATGTCGATGGAGATCTCTTAACAATGCT----- 405  
QY 223 GlyGlyAspGlyAsnGlyGlyAsnCysPheProLeuThrValProLeuThrIle 242  
DB 406 -----GGCCAAATGAGTTACCATGAGCTACTCATCATCATTCATCAA 447  
QY 243 AsnGlnSerClnGlnLysArgAspValGlyLeuSerGlyGlyGluAspValGlyAspAsn 262  
DB 448 AATCATAGCTCAAAATGTCATCGATTTGTGCCCATCTGATCAAAACCTCAATTTCAA 507  
QY 263 ValTyrProValArgMetThrValPheIleAsnGluMetProIleGluValValSerGly 282  
DB 508 TACCAACAAGGGGTATACGGTGTATTAACGGAGTTCCGACAGAGTGCAGAGGA 567  
QY 283 LeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGln 302  
DB 568 GGAATAGACATGAAGCAACGTTTGGAGAGATTTGGTTTGGTGATTCCTCAGGTGTT 627  
QY 303 ProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyr 322  
DB 628 CTTCTCTCTACTGATGAGTTGGTTTTTGTGTCATAGCTTACACATGTTGAGCTTAT 687  
QY 323 TyrLeuIle 325  
DB 688 TTCCTGGTT 696  
RESULT 7  
AAS02222  
ID AAS02222 standard; cDNA; 1367 BP.  
XX AAS02222;  
AC AAS02222;  
XX 16-JUL-2001 (first entry)  
XX Corn Muschel (WUS) cDNA from clone p0058\_chpab57r.  
DE Wuschel; WUS; apical meristem; organ formation; homeodomain protein; ss;  
XX transcriptional regulator; proliferation; corn; soybean; wheat; rice;  
KW alfalfa; sunflower; canola; cotton; enzyme activity; Genetic marker;  
XX plant transformation; transgenic plant; p0058\_chpab57r.  
OS Zea mays.  
XX Key Location/Qualifiers  
FH 202..963  
CDS /\*tag= a  
FT /product= "Corn WUSCHEL protein"  
XX WO200123575-A2.  
XX 05-APR-2001.  
XX 28-SEP-2000; 2000WO-US026648.  
XX 30-SEP-1999; 99US-0157216P.  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX Cahoon RE, Lowe K;  
XX WPI; 2001-258137/26.  
DR P-PSDB; AAU01525.

XX Novel nucleic acid fragments encoding WUS proteins useful for transiently  
 PT modulating WUS protein level in plant cells, as probes for genetically  
 PT and physically mapping WUS genes and as markers.

XX Claim 4; Page 49-50; 61pp; English.

XX The sequence represents cDNA from a clone, which encodes a Wuschel (WUS)  
 CC protein. The WUS protein plays a key role in initiation and maintenance  
 CC of the apical meristem of a plant which contains stem cells, allowing for  
 CC organ formation. WUS encodes a homeodomain protein which is thought to  
 CC function as a transcriptional regulator. The DNA or its complement is  
 CC useful for transiently modulating the level of WUS protein in a plant  
 CC cell and meristem proliferation can be induced through its addition. This  
 CC involves transforming a plant with WUS DNA to produce a transformed  
 CC meristem which is grown into a regenerated plant (corn, soybean, wheat,  
 CC rice, alfalfa, sunflower, canola or cotton). The protein and DNA are  
 CC useful in the identification of sequences that affect WUS level or enzyme  
 CC activity, or that encode homologous proteins from the same or other plant  
 CC species. The polypeptides can then be used as probes to map the genes that  
 CC fragments of the DNA may be used as probes to map the genes that they are  
 CC part of and as markers for traits linked to those genes, which is useful  
 CC in plant breeding to develop lines with desired phenotypes. The nucleic  
 CC acid sequences are useful to create transgenic plants in which the WUS  
 CC polypeptides are present in higher or lower levels than normal or in cell  
 CC types or developmental stages in which they are not normally found

XX Sequence 1367 BP; 291 A; 431 C; 394 G; 251 T; 0 U; 0 Other;

Alignment Scores: Length: 1367  
 Pred. No.: 1.37e-25  
 Score: 328.50  
 Percent Similarity: 45.00%  
 Best Local Similarity: 31.33%  
 Query Match: 18.76%  
 DB: 4  
 Gaps: 7

US\_09-787-737-2 (1-325) x AAS02222 (1-1367)

QY 35 TyrSerAsnCys-AsnLeuSerSerPheSerPheSerArgIleProAspProLysPr 54  
 DB 186 TATCATACCTGTGCATGAGGAGGCGACTGAGCCGCGGCGGCGGAGGCC 245  
 QY 54 O-----ArgTIPAsnProLysProGluGlnIleArgIleLeuSerIlePheAs 71  
 DB 246 GTTCGGTTCGGGTGGAGCGCCAGCGGAGCGAGATCACTCATCTCGAGTCCATCTCAA 305  
 QY 71 nSerGlyThrIleAsnProProArgGluGluIleGlnArgIleArgIleGlnGlu 91  
 DB 306 CAGCGGCGATGTTGAACCGCGCCCAAGGAGCGAGCGTCCGATCCGAGCTCTGGAGCG 365  
 QY 91 uTyrGlyGlnIleGlyAspAlaAsnValPheTyrTriPheGlnAsnArgLysSerArgAl 111  
 DB 366 CTTCCGGCGCGGTGGCGGAGCGCAACGCTTCTACTGTTCCAGAACCGCGCTCCCGCTC 425  
 QY 111 aLysHisIysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrVa 131  
 DB 426 CCGCGGCGCGGCGAGCGCCAGCTGCAGCGGCGGCGGCGCTCTCTCTCGTCGGGATC 485  
 QY 131 lileProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTy 151  
 DB 486 GCCCGCCACG-----ACGCGCTCGC 506  
 QY 151 rProValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHi 171  
 DB 507 ACCG-----GGACAGCGCGCGCTTCGTCGAGCGGCGGATGTTCCGCGCAGCGCGC 557  
 QY 171 sAsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValProValVa 191  
 DB 558 CACCTACGCGCTCGTCGGCGTCGCGCTCTGG-----CCGCGCGCGCGC 599  
 QY 191 lThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGln 211

Db 600 GTCGTGCGAGGGGATGATGGCGGACCTGGACTACGGCGGC----- 639  
 QY 211 uAsnPhAspLysIleProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGlyGlyAs 231  
 Db 640 ---GGCGACACCTGTTCCGCTATCTCGGCGACATGGGCTACGCCAGCGGCGTGGCTC 695  
 QY 231 nCysPheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspVa 251  
 Db 696 CGGTCGCGCTCTCGGCGCGCTCGCCACCGACGACGACGACGACGACGACGACGACG 744  
 QY 251 lGlyLeuSerGlyGlyGluAspValGlyAspAsnValTyr-----ProVa 266  
 Db 745 -----CTTACTACTCGCGTGCACCCAGC 770  
 QY 266 lArgMetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVa 286  
 Db 771 GAGCATGACGGTGTTCATCATGCGGTGGCGACGAGGTGCGCGGGGCGGCGATCGACCT 830  
 QY 286 lIysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuTh 306  
 Db 831 GCGGTCCATGTTCCGCGGAGGACGTGCTGTGTGCACTCCACCGCGGCTCTCTCCCGT 890  
 QY 306 rAspGluPheGlyValThrTyrGlnProIleGlnAsnGlyAlaIleTyrTyrLeuIle 325  
 Db 891 CAACGATACGCGCTGCTCAOCGACAGGCTGCAGATGGCGGAGAGCTACTTCTGCTC 948  
 RESULT 8  
 ADD30299  
 ID ADD30299 standard; cDNA; 816 BP.  
 XX  
 AC ADD30299;  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Plant yield-related polynucleotide clone G1587.  
 XX  
 KW ds; transcription factor; transgenic plant; growth rate; senescence;  
 KW seed germination rate; plant vigor; seedling vigor.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO2003013227-A2.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 09-AUG-2002; 2002WO-US025805.  
 XX  
 PR 09-AUG-2001; 2001US-0310847P.  
 PR 19-NOV-2001; 2001US-0336049P.  
 PR 11-DEC-2001; 2001US-0338692P.  
 PR 14-JUN-2002; 2002US-00171468.  
 XX  
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 XX  
 PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;  
 PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;  
 PI Broun PE;  
 XX  
 XX WPI; 2003-248221/24.  
 DR P-ESDB; ADD30300.  
 XX  
 PT New plant transcription factor polynucleotides and polypeptides, useful  
 PT in producing transgenic plants with commercially valuable properties,  
 PT such as an alteration in a plant growth characteristic, e.g. growth rate  
 PT or apomixis.  
 XX  
 PS Disclosure; SEQ ID NO 328; 454pp; English.  
 XX  
 CC The invention relates to a number of isolated Arabidopsis thaliana cDNA  
 CC sequences and their encoded proteins which are especially transcription  
 CC factor related cDNA's and proteins. The isolated or recombinant plant  
 CC transcription factor polynucleotides and polypeptides are useful in  
 CC producing transgenic plants with commercially valuable properties, i.e.





us-09-787-737-2.rng

Mon Sep 13 10:30:57 2004

```
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144322P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149369P.
PR 17-AUG-1999; 99US-0149175P.

PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Alignment Scores:  
Pred. No.: 5,74e-13  
Score: 212.50  
Percent Similarity: 39.78%  
Best Local Similarity: 28.25%  
Query Match: 12.14%  
DB: 3  
Length: 1326  
Matches: 76  
Conservative: 31  
Mismatch: 87  
Indels: 75  
Gaps: 12

US-09-787-737-2 (1-325) x AAC38580 (1-1326)

Qy 18 AsnAsnHisHisHisGln-----HisGlnLeaspthr 29  
Db 283 AACGGCAATGATCACCATCATGCTATATACCCGATCATCAGGACACGATATTCATGAA 342

Qy	30	ProSerTyrMethIstYrSerAsnCysAsnLeuSerSerPheSerSerAspArgIle	49
Db	343	-----CGGAAACAACATT	357
Qy	50	ProAspProLysPro-----ArgTTPAsnProLysProGluGlnIleArgIleLeuGluSer	68
Db	358		
Qy	69	IlePheAsnSerGlyThrIleAsnProProArgGluGluIleGlnArgIleArgIleArg	88
Db	418		
Qy	89	LeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLys	108
Db	478		
Qy	109	SerArgAlaIlyshIstLysLeuArgValHisLysSerProLysMetSerLysLysAsp	128
Db	538		
Qy	129	LysThrValIleProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThr	148
Db	595	AAAGCGTCAAGATTCAATCAGCGTGTCTCGAGTT-----GATCAGACA	642
Qy	149	GlyLeuTyrProValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPhe	168
Db	643	AAAGCTCTGCCA-----TCCTTTT	660
Qy	169	ProValHisAsnAspProSerAlaLaGlnSerAlaPheGlyPheGlyAspPheValVal	188
Db	661	CCACACACAAACCGACCACAG-----	681
Qy	189	ProValValThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGlu	208
Db	682	CCACAGCATGAATTAGATCCTCGAGTTTACATAAAGACAACAATGCTAATAATGAAGAT	741
Qy	209	-----ThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrClyGlyAsp	225
Db	742	CATGGGACGACTGAAGATCTGATCAGAGGGCATCAGAGTTGGTAAATACGCCACATGG	801
Qy	226	GlyAsnGlyGlyAsnCysPheProProLeuThrValProLeuThrIleAsnGlnSer	245
Db	802	AGAAAT-----CTTGTTACTTGGTCGATACTCAACAA	834
Qy	246	GlnGluLys-----ArgAspValClyLeuSerGly-----GlyGluAspValGlyAsp	261
Db	835	CCGAGAGAGATTATATTCGAGAAATATCAACGGAGAGAAGAAACGAGGACAAC	894
Qy	262	-----AsnValTyrProValArg	267
Db	895	CGGACTTTAAATCTCTTTCGGTAGG	921
RESULT 10			
AC	AA47056		
ID	AA47056	standard; DNA; 1209 Bp.	
AC	AA47056;		
DT	18-OCT-2000	(first entry)	
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 52401.	
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
XX			

us-09-787-737-2.rng

Db 477 ATGGGGCTAACGATCATATCATCTCTACTTACCACATCATCAGGTGTTCCTCATGCG 536  
 Qy 126 LysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPheValAsn 145  
 Db 537 AGACCTGTATTCGGTCAAGTAACTTAACCAAGACCATCATCTCTATCATCATTAAC 596  
 Qy 146 GlnGluThrGlyLeuTyProValGlnAsnAsnGluLeuValThrGluProAlaGly 165  
 Db 597 AAGCCA-----TATCCGAGCTTCAATAACGGGAATTAAATCATGCAAGCTCAGT 647  
 Qy 166 PheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAsp 185  
 Db 648 ACTGAATGTGGTGTGTAATGCTTCTAATGGCTTACATGAGTAGCCATGTCTATGGATCT 707  
 Qy 186 PheValValProValValThrGluGluGlyMetAlaPheSerThrValAsnAsnGly--- 204  
 Db 708 -----ATGGAACAAGACTGTTCTATGAATTACCAACACGATGAGTGAGGATGG 755  
 Qy 205 ValAsnLeuGluThrAsnGluAsn-----PheAspLysIleProAla 218  
 Db 756 GCAACATGATCATCATCTACTCATCTGACCTTACAACTTCTTCGATAGA-----GCA 809  
 Qy 219 IleAsnLeuTyGlyGlyAspGlyAsnGly-----GlyGlyAsnCysPhe 233  
 Db 810 AAGCCTCTGTTGGTCTAGAGTCTCATCAAGAGAGAGAGATGTGGTGGCGATGCTTAT 869  
 RESULT 11  
 ID ADD30291 standard; cDNA; 1227 BP.  
 AC ADD30291;  
 XX  
 DT 15-JAN-2004 (first entry)  
 DE Plant yield-related polynucleotide clone G1540.  
 KW ds; transcription factor; transgenic plant; growth rate; senescence;  
 KW seed germination rate; plant vigor; seedling vigor.  
 XX Arabidopsis thaliana.  
 XX WO2003013227-A2.  
 PN 20-FEB-2003.  
 PD 09-AUG-2002; 2002WO-US025805.  
 PF 09-AUG-2001; 2001US-0310847P.  
 PR 19-NOV-2001; 2001US-0336049P.  
 PR 11-DEC-2001; 2001US-0338692P.  
 PR 14-JUN-2002; 2002US-00171468.  
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;  
 PI Pilgram ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;  
 PI Broun PE;  
 XX WPI; 2003-248221/24.  
 DR P-FSDB; ADD30292.  
 XX  
 PT New plant transcription factor polynucleotides and polypeptides, useful  
 PT in producing transgenic plants with commercially valuable properties,  
 PT such as an alteration in a plant growth characteristic, e.g. growth rate  
 PT or apomixis.  
 PS Disclosure; SEQ ID NO 320; 454pp; English.  
 XX  
 CC The invention relates to a number of isolated Arabidopsis thaliana cDNA  
 CC sequences and their encoded proteins which are especially transcription  
 CC factor related cDNA's and proteins. The isolated or recombinant plant  
 CC transcription factor polynucleotides and polypeptides are useful in  
 CC producing transgenic plants with commercially valuable properties, i.e.

CC modified or altered desirable traits as compared to a reference plant,  
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,  
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and  
 CC flower senescence. Sequence information related to the polynucleotides  
 CC and polypeptides can also be used in bioinformatic search methods. The  
 CC transgenic plant is useful for growing a progeny plant from a parent  
 CC plant. This sequence represents one of the cDNAs of the invention.  
 XX  
 SQ Sequence 1227 BP; 351 A; 292 C; 239 G; 345 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1,34e-11 Length: 1227  
 Score: 199.50 Matches: 67  
 Percent Similarity: 41.54% Conservative: 41  
 Best Local Similarity: 25.77% Mismatches: 97  
 Query Match: 11.39% Indels: 55  
 DB: 9 Gaps: 10

US-09-787-737-2 (1-325) x ADD30291 (1-1227)

Qy 15 LysProCysAsnAsnAsnHisHisGlnHisGluIleAspThrProSerTyrMetHis 34  
 Db 125 GAGCCGCCACAGCATCAGCATCATCAT---CATCAGCCGACCAAGAAAGC----- 172  
 Qy 35 TyrSerAsnCysAsnLeuSerSerPheSerSerAspArgIleProAspProLysPro 54  
 Db 173 ---GGCAACAACAACAAGTCCGGCTCTGTGTGTACAGGTGTCCGACGACGACGACG 229  
 Qy 55 ArgTrpAsnProLysProGluGlnIleArgIleLeuGluSerIle---PheAsnSerGly 73  
 Db 230 AGGTGACACACGACGACGAGCAATCAAAATCCCTCAAGAACTTTTACTACCAATGCA 289  
 Qy 74 ThrIleAsnProArgGluGluIleGlnArgIleArgIleArdGluGlnGluTyGly 93  
 Db 290 ATCCGCTCACCAACAGCCGATCAGATCCAGAGATCATGCAAGGTGAGACAGTTCGA 349  
 Qy 94 GlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHis 113  
 Db 350 AAGATTGAGGCAAGAACGCTCTTTTACTGGTCCAGAACCATTAAGGCTGTGAGCGTCAG 409  
 Qy 114 LysLeuArgVal----- 117  
 Db 410 AAGAAGATTCACCGGAACAACATGACCAACCATCTTCATCACCCCAACTCGGTATG 469  
 Qy 118 -----HisHisLysSerProLysMetSer 125  
 Db 470 ATGGCGGCTAACGATCATTCATCTCTACTTCCACATCATCAGGTGTTCCCATGCGAG 529  
 Qy 126 LysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPheValAsn 145  
 Db 530 AGACCTGCTAATTCGGTCAACGTTAAACTTAACCAAGACCATCATCTCTATCATCATTAAC 589  
 Qy 146 GlnGluThrGlyLeuTyProValGlnAsnAsnGluLeuValThrGluProAlaGly 165  
 Db 590 AAGCCA-----TATCCGAGCTTCAATAACGGGAATTTAAATCATGCAAGCTCAGT 640  
 Qy 166 PheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAsp 185  
 Db 641 ACTGAATGTGGTGTGTAATGCTTCTAATGGCTTACATGAGTAGCCATGTCTATGGATCT 700  
 Qy 186 PheValValProValValThrGluGluGlyMetAlaPheSerThrValAsnAsnGly--- 204  
 Db 701 -----ATGGAACAAGACTGTTCTATGAATTACCAACACGATGAGTGAGGATGG 748  
 Qy 205 ValAsnLeuGluThrAsnGluAsn-----PheAspLysIleProAla 218  
 Db 749 GCAACATGATCATCATCTACTTCGACCTTACAACTTCTTCGATAGA-----GCA 802  
 Qy 219 IleAsnLeuTyGlyGlyAspGlyAsnGly-----GlyGlyAsnCysPhe 233  
 Db 803 AAGCCTCTGTTGGTCTAGAGTCTCATCAAGAGAGAGATGTGGTGGCGATGCTTAT 862

RESULT 12

Db	161	ATTCAAGGCAAGAATGTGTCTACTGGTTTCAGAACCAAAAGCCAGAGAAAGCAGAG 220
Qy	115	LeuArg-----ValHisHisLysSerProLysMetSerIysLysAspLysThr 130
Db	221	CAAAAGCGTTAACAGCAGCATGCACCGAGTAGCTGCTACTCAGCAAAAGAACTCCAACA 280
Qy	131	ValIle 132
Db	281	ACAATA 286
RESULT 13		
AAS02219		
ID	AAS02219	standard; cDNA; 1338 BP.
XX		
AC	AAS02219;	
XX		
DT	16-JUN-2001	(first entry)
XX		
DE	Corn Wuschel (WUS) cDNA from clone cpilc_pk012_p19.	
XX		
KW	Wuschel; WUS; apical meristem; organ formation; homeodomain protein; ss;	
KW	transcriptional regulator; proliferation; corn; soybean; wheat; rice;	
KW	alfalfa; sunflower; canola; cotton; enzyme activity; genetic marker;	
KW	plant transformation; transgenic plant; cpilc_pk012_p19.	
XX		
OS	Zea mays.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	238..900
FT		/*tag= a
FT		/product= "Corn WUSCHEL protein"
XX		
PN	WO200123575-A2.	
XX		
PD	05-APR-2001.	
XX		
PF	28-SEP-2000; 200WO-US026648.	
XX		
PR	30-SEP-1999; 99US-0157216P.	
XX		
PA	(DUPO ) DU PONT DE NEMOURS & CO E I.	
XX		
PI	Cahoon RE, Lowe K;	
XX		
DR	WPI; 2001-258137/26.	
XX		
DR	P-PSDB; AAU01522.	
XX		
PT	Novel nucleic acid fragments encoding WUS proteins useful for transiently	
PT	modulating WUS protein level in plant cells, as probes for genetically	
PT	and physically mapping WUS genes and as markers.	
XX		
PS	Claim 4; Page 46; 61pp; English.	
XX		
CC	The sequence represents cDNA from a clone, which encodes a Wuschel (WUS)	
CC	protein. The WUS protein plays a key role in initiation and maintenance	
CC	of the apical meristem of a plant which contains stem cells, allowing for	
CC	organ formation. WUS encodes a homeodomain protein which is thought to	
CC	function as a transcriptional regulator. The DNA or its complement is	
CC	useful for transiently modulating the level of WUS protein in a plant	
CC	cell and meristem proliferation can be induced through its addition. This	
CC	involves transforming a plant with WUS DNA to produce a transformed	
CC	meristem which is grown into a regenerated plant (corn, soybean, wheat,	
CC	rice, alfalfa, sunflower, canola or cotton). The protein and DNA are	
CC	useful in the identification of sequences that affect WUS level or enzyme	
CC	activity, or that encode homologous proteins from the same or other plant	
CC	species. The polypeptides can then be used to prepare antibodies.	
CC	Fragments of the DNA may be used as probes to map the genes that they are	
CC	part of and as markers for traits linked to those genes, which is useful	
CC	in plant breeding to develop lines with desired phenotypes. The nucleic	
CC	acid sequences are useful to create transgenic plants in which the WUS	
CC	polypeptides are present in higher or lower levels than normal or in cell	
CC	types or developmental stages in which they are not normally found	

```

XX SQ Sequence 1338 BP; 328 A; 369 C; 389 G; 252 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.08e-10 Length: 1338
Score: 187.50 Matches: 45
Percent Similarity: 45.89% Conservative: 22
Best Local Similarity: 30.82% Mismatches: 50
Query Match: 10.71% Indels: 29
DB: 4 Gaps: 5

US-09-787-737-2 (1-325) x AAS02219 (1-1338)
QY 25 HisGluIleAspThrProSerTyrMetHisTyrSerAsn-----CysAsnLeuSerSer 42
Db 124 CACACATATAGCTGAAGCAATATATCCACTTCGTTAACTGGCGGTAGTGTAGCTGCG 183
QY 43 -----SerPheSer----- 45
Db 184 ATCGCTGCAAACTACAGGGGTGTAGTGTATCGATCGCTGATCATATATACCATGGAG 243
QY 46 -----SerAspArgIleProAspProLysProArgTyrAsnProLysProGluGlnIle 63
Db 244 GCGCTAGCGGGCGGTAGCGTCAAGTGGCGGGGTGAACCTACGGCGGAGCAGGTG 303
QY 64 ArgIleLeuGluSerIlePheAsnSerglyThrIleAsnProProArgGluGluIleGln 83
Db 304 AAGGTCTGACGGAGCTCTTCGCGCGGGCTCGGACGCCAGCCAGCAGGAGCATCCAG 363
QY 84 ArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrp 103
Db 364 CGCATCTCCACCACCTTCGCGCGCTTCGCAAGTGGAGAGCAAGACGTCTTCTACTGG 423
QY 104 PheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLys 123
Db 424 TTCAGAACCAACAGGCCCGCGAGCGCCAC-----CACCAAGAGAGCGCCGC 471
QY 124 MetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPhe 143
Db 472 CGCGCGCGCTCGTGCCTC-----TCCCGCAGCGCGGAGCGGAGCGGAGC 519
QY 144 ValAsnGlnThrGly 149
Db 520 AACACGAGGAGACGGC 537

RESULT 14
AAS02227
ID AAS02227 standard; cDNA; 741 BP.
AC AAS02227;
DT 16-JUL-2001 (first entry)
DE Soybean Wuschel (WUS) cDNA from clone ses4d_pk0033_c8.
KW Wuschel; WUS; apical meristem; organ formation; homeodomain protein; ss;
KW transcriptional regulator; proliferation; corn; soybean; wheat; rice;
KW alfalfa; sunflower; canola; cotton; enzyme activity; genetic marker;
KW plant transformation; transgenic plant; ses4d_pk0033_c8.
OS Glycine max.
XX FH Key Location/Qualifiers
XX CDS 27..665
XX FT /*tag= a
XX FT /product= "Corn WUSCHEL protein"
XX WO200123575-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US026648.
XX

```

30-SEP-1999; 99US-0157216P.

(DUPO) DU PONT DE NEMOURS & CO E I.

Cahoon RE, Lowe K;

WPI; 2001-258137/26.

P-PSDB; AAU01530.

Novel nucleic acid fragments encoding WUS proteins useful for transiently modulating WUS protein level in plant cells, as probes for genetically and physically mapping WUS genes and as markers.

Claim 4; Page 56; 6lpp; English.

The sequence represents cDNA from a clone, which encodes a Wuschel (WUS) protein. The WUS protein plays a key role in initiation and maintenance of the apical meristem of a plant which contains stem cells, allowing for organ formation. WUS encodes a homeodomain protein which is thought to function as a transcriptional regulator. The DNA or its complement is useful for transiently modulating the level of WUS protein in a plant cell and meristem proliferation can be induced through its addition. This involves transforming a plant with WUS DNA to produce a transformed. This meristem which is grown into a regenerated plant (corn, soybean, wheat, rice, alfalfa, sunflower, canola or cotton). The protein and DNA are useful in the identification of sequences that affect WUS level or enzyme activity, or that encode homologous proteins from the same or other plant species. The polypeptides can then be used to prepare antibodies. Fragments of the DNA may be used as probes to map the genes that they are part of and as markers for traits linked to those genes, which is useful in plant breeding to develop lines with desired phenotypes. The nucleic acid sequences are useful to create transgenic plants in which the WUS polypeptides are present in higher or lower levels than normal or in cell types or developmental stages in which they are not normally found.

Sequence 741 BP; 238 A; 155 C; 156 G; 192 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.51e-10	Length:	741
Score:	187.00	Matches:	62
Percent Similarity:	38.94%	Conservative:	19
Best Local Similarity:	29.81%	Mismatches:	55
Query Match:	10.68%	Indels:	72
DB:	4	Gaps:	9

US-09-787-737-2 (1-325) x AAS02227 (1-741)

QY 43 SerPheSerSerAsp-----ArgIleProAspProLysProArgTrp 56

Db 33 AGTCACATAGTAGTGTGAGCGGAGAGATGTAGGACTCATTCAGTTTCAAGGTGG 92

QY 57 AsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerglyThrIleAsn 76

Db 93 AGTCTCAAAAGGACCAATAGACATGTTAGAGAACCTTTACACGAGGGAATAGGACT 152

QY 77 ProProArgGluGluIleGlnArgIleArgLeuGlnGluTyrGlyGlnIleGly 96

Db 153 CCCAGCAGCTGACCAATACACAGATTACCTCTAGGCTCAGGCTTATGGTCACATCGAG 212

QY 97 AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116

Db 213 GGAAGAAATGTCTTCTACTGTGTTTCAAAATCACAAGCTCGTCAAGACAGAGCTG--- 269

QY 117 ValHisHisLysSerProLysMetSerLysLysAspLysThrValIleProSerThr--- 135

Db 270 -----ATGAAGCAACACCATTCGATATTCGAATCGC 302

QY 136 -----AspAlaAspHis-----CysPheGlyPheValAsnGln 146

Db 303 TTTCTCTGCTGCCCTCCACCCCATTTGCCAAATGTTGCTCGCTCCATATTTGTTGAA 362

QY 147 GluThrGly-----LeuTyrProValGlnAsnAsnGluLeuVal----- 159

CC useful in the identification of sequences that affect WUS level or enzyme  
CC activity, or that encode homologous proteins from the same or other plant  
CC species. The polypeptides can then be used to prepare antibodies.  
CC Fragments of the DNA may be used as probes to map the genes that they are  
CC part of and as markers for traits linked to those genes, which is useful  
CC in plant breeding to develop lines with desired phenotypes. The nucleic  
CC acid sequences are useful to create transgenic plants in which the WUS  
CC polypeptides are present in higher or lower levels than normal or in cell  
CC types or developmental stages in which they are not normally found  
XX  
SQ Sequence 844 BP; 292 A; 163 C; 166 G; 223 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3 25e-09 Length: 844  
Score: 175.50 Matches: 43  
Percent Similarity: 51.64% Conservative: 20  
Best Local Similarity: 35.25% Mismatches: 34  
Query Match: 10.02% Indels: 25  
DB: 4 Gaps: 5

US-09-787-737-2 (1-325) x AAS02226 (1-844)

QY 55 ArgTrrAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThr 74  
Db 74 CGATGGAGTCTTACAAAGGAGCAATAGACATGTTGGAGAACTTTTACAAAGCAGGAATA 133  
QY 75 IleAsnProArgGluGluIleGlnArgIleArgIleLeuGluGlnGlyGln 94  
Db 134 AGGACTCCAGCAGCTGAGCAATACACAGATTACTCTAGGCTTAGGCTTATGGTTAC 193  
QY 95 IleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLys 114  
Db 194 ATCGAGGGAAAAATGCTCTTCTACTGTTTCAAAATCACAAGCGCGCCAAAGACAGAAG 253  
QY 115 LeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIleProSer 134  
Db 254 CTCAAG-----CAGAGCAACAAGACATTGCA----- 280  
QY 135 ThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrProValGln 154  
Db 281 -----TACTGCAATTGCTTCTTCATGCGCTCC-----CACCCCATTTGC 319  
QY 155 AsnAsnGluLeuValThrGluPro-----AlaGlyPheLeuPhePro 169  
Db 320 CAAAT-----GTTCTCTGCGCTCCATATGTTTGCAAAAGAGTGGATTGAGTTTAT 373  
QY 170 ValHis 171  
Db 374 CCTCAC 379

Search completed: September 10, 2004, 09:26:26  
Job time : 485 secs

Db 363 CGGAGTGATTACAGCTTTATCTTCAACATCGAAGTGGCTTGCAGTGGAGTTAAGT 422  
QY 160 ValThrGluProAlaGlyPhe----- 166  
Db 423 TCACTGGCCCTTAGGCATGCAAGAGATGTTGATGGCATGCGAGTAGTGAACACCG 482  
QY 167 -----LeuPheProVal----- 170  
Db 483 GATTGTAAACCGTGAAGTCTTAACCTCTTCTCTTCATCCAAACCGGCATTTTGAAGAA 542  
QY 171 -----HisAsnAspProSerAlaGlnSerAlaPheGlyPheGlyAspPheVal 187  
Db 543 AAAACAATCTCAAGTGCCTTCCCTGCTTCACTTCT----- 581  
QY 188 ValProValThrGluGluGly 195  
Db 582 GTTGTGCTGTTGATGAAGATGGT 605

RESULT 15  
AAS02226  
ID AAS02226 standard; cDNA; 844 BP.  
XX  
AC AAS02226;  
XX  
DT 16-JUL-2001 (first entry)  
XX  
DE Soybean Wuschel (WUS) cDNA from clone scrlc\_pk001\_d2.  
XX  
KW Wuschel; WUS; apical meristem; organ formation; homeodomain protein; ss;  
KW transcriptional regulator; proliferation; corn; soybean; wheat; rice;  
KW alfalfa; sunflower; canola; cotton; enzyme activity; genetic marker;  
KW plant transformation; transgenic plant; scrlc\_pk001\_d2.  
XX  
OS Glycine max.  
XX

Key Location/Qualifiers  
FH 8.661  
FT CDS  
FT /\*tag= a  
FT /product= "Corn WUSCHEL protein"  
FT /partial  
FT /note= "No start codon"

XX WO200123575-A2.  
XX  
XX 05-APR-2001.  
XX  
XX 28-SEP-2000; 2000WO-US026648.  
XX  
XX 30-SEP-1999; 99US-0157216P.  
XX  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
XX Cahoon RE, Lowe K;  
XX  
XX WPI; 2001-258137/26.  
XX P-PSDB; AAU01529.  
XX

Novel nucleic acid fragments encoding WUS proteins useful for transiently  
PT modulating WUS protein level in plant cells, as probes for genetically  
PT and physically mapping WUS genes and as markers.  
XX  
XX Claim 4; Page 55; 61pp; English.

XX The sequence represents cDNA from a clone, which encodes a Wuschel (WUS)  
XX protein. The WUS protein plays a key role in initiation and maintenance  
XX of the apical meristem of a plant which contains stem cells, allowing for  
XX organ formation. WUS encodes a homeodomain protein which is thought to  
XX function as a transcriptional regulator. The DNA or its complement is  
XX useful for transiently modulating the level of WUS protein in a plant  
XX cell and meristem proliferation can be induced through its addition. This  
XX involves transforming a plant with WUS DNA to produce a transformed  
XX meristem which is grown into a regenerated plant (corn, soybean, wheat,  
XX rice, alfalfa, sunflower, canola or cotton). The protein and DNA are



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 10, 2004, 09:01:29 ; Search time 91 Seconds  
(without alignments)  
1981.967 Million cell updates/sec

Title: US-09-787-737-2  
Perfect score: 1751  
Sequence: 1 MSSNNKWPMSFKSPKNN.....TDFGVGYQLQNGAIYILI 325

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09787737 @CGN 1.1.69 @runat 01092004.112456.23620 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	6.7	4791	4	US-09-328-352-1416
2	104	5.9	3602	4	US-09-402-929-1
3	104	5.9	6775	4	US-09-402-929-4
4	103.5	5.9	860	4	US-09-031-962D-3
5	102.5	5.9	1404	4	US-09-489-039A-619
6	101.5	5.8	12286	4	US-09-322-478-17
7	98.5	5.6	2992	1	US-08-426-236-3
8	96.5	5.5	29604	3	US-08-781-891-207
9	96.5	5.5	29604	4	US-09-618-166-207
10	96.5	5.5	4403765	3	US-09-103-840A-2
11	96.5	5.5	4411529	3	US-09-103-840A-1
12	96	5.5	9829	4	US-09-322-478-19

C	13	93.5	5.3	2260	4	US-09-889-463A-35	Sequence 35, Appl
	14	93	5.3	864	4	US-09-031-962D-1	Sequence 1, Appli
	15	92.5	5.3	936	4	US-09-252-991A-14226	Sequence 14226, A
	16	91	5.2	1318	1	US-08-485-618-100	Sequence 100, App
	17	91	5.2	1318	1	US-08-605-672-100	Sequence 100, App
	18	91	5.2	1318	2	US-08-482-293A-100	Sequence 100, App
	19	91	5.2	1318	2	US-08-943-363-100	Sequence 100, App
	20	91	5.2	1318	3	US-09-193-043-100	Sequence 100, App
	21	91	5.2	1318	4	US-09-688-307A-100	Sequence 100, App
	22	91	5.2	1318	4	US-09-350-259-100	Sequence 100, App
	23	91	5.2	2329	3	US-08-327-219-80	Sequence 100, App
	24	91	5.2	2328	3	US-08-327-219-1	Sequence 80, Appl
	25	91	5.2	3238	3	US-08-927-219-3	Sequence 1, Appli
	26	90.5	5.2	1065	3	US-08-591-685-6	Sequence 3, Appli
	27	90	5.1	2293	1	US-08-604-913B-12	Sequence 6, Appli
	28	90	5.1	3004	1	US-08-276-213-6	Sequence 12, Appl
	29	89.5	5.1	5362	2	US-08-853-310-3	Sequence 6, Appli
	30	89.5	5.1	10480	4	US-09-732-615-13	Sequence 3, Appli
	31	89	5.1	2679	1	US-08-479-328-1	Sequence 13, Appl
	32	89	5.1	2679	1	US-08-761-119-1	Sequence 1, Appli
	33	89	5.1	2679	2	US-08-668-128B-1	Sequence 1, Appli
	34	89	5.1	2679	2	US-08-905-445-1	Sequence 1, Appli
	35	89	5.1	2679	3	US-08-959-625-1	Sequence 1, Appli
	36	89	5.1	2679	3	US-09-008-466-1	Sequence 1, Appli
	37	89	5.1	2679	3	US-08-580-980A-1	Sequence 1, Appli
	38	89	5.1	2679	3	US-09-053-453-1	Sequence 1, Appli
	39	89	5.1	2679	3	US-08-644-116A-1	Sequence 1, Appli
	40	89	5.1	2679	4	US-09-437-858-1	Sequence 1, Appli
	41	89	5.1	4200	1	US-07-841-654B-1	Sequence 1, Appli
	42	89	5.1	4200	1	US-07-946-234A-1	Sequence 1, Appli
	43	89	5.1	4200	1	US-08-123-161A-1	Sequence 1, Appli
	44	89	5.1	4200	1	US-08-483-278-1	Sequence 1, Appli
	45	89	5.1	4200	5	PCT-US93-01560-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-328-352-1416  
; Sequence 1416, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 1416  
; LENGTH: 4791  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-1416

Alignment Scores:

Pred. NO.: 0.00655 Length: 4791  
Score: 117.00 Matches: 83  
Percent Similarity: 32.90% Conservative: 44  
Best Local Similarity: 21.50% Mismatches: 139  
Query Match: 6.68% Indels: 120  
DB: 4 Gaps: 20

US-09-787-737-2 (1-325) x US-09-328-352-1416 (1-4791)

QY	8	TriProSerMetPhelySerLysProCysAsnAsnHisHisGlnHisGlu---	26
Db	3664	TGGGAGTCT---TTTAAATCTGCACAGCAATATATCAAGCACTATTTATGAACCT	3720
QY	27	-----	-----IleAspThr 29
Db	3721	GATAGCTTTGTTCCACTATTACAGCAGGTTATAGGATTTTATCAACTGATGAACCT	3780

Mon Sep 13 10:30:58 2004

us-09-787-737-2.rni

```

QY 30 ProSerTyrMetHisTyrSer-----Asn 37
Db 3781 CCTGATTATCAGGAATATCAACAAACCTTATTCTATTATTATAAGACCCAGTAGTAAT 3840
QY 38 CysAsnLeuSerSerSer-----PheSerSerAspArg 48
Db 3841 CGTAATCTTGTAAGAGCGTACTGCTTTAGAGAGATTTCATCTACCACTGTAUCAG 3900
QY 49 IleProAspProLysProArgTyrAsnProLysProGluGlnIle-----ArgIleLeuGlu 67
Db 3901 GTCGAACACACAAACCATGACCAATATCCGAGGAGAGTGTGTATGGAAATTTCTCAA 3960
QY 68 SerIlePheAsnSer-----GlyThrIleAsnProProArgGlu 80
Db 3961 GATACATGGCGCGGTGTCGCAAAATTAAGCACTAAATCAAGATAATCCA----- 4011
QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100
Db 4012 ---TTTGAACAAATAATTTACGTTTCAA-----GGCCAGTATTACGCGTGAACA 4062
QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLys 120
Db 4063 GAACGCACTATAACGGTTAUGTTATTATGAACTCATAGCGCAAGATATGTAAGTAAA 4122
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
Db 4123 GACCCG----- 4128
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValVal 160
Db 4129 -----ATTGGTTAGAGGGTGTATGAATACCTCTAGTTATGTA 4167
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaGlnSerAla 180
Db 4168 AGTGATCCAAATCAATGGATT-----GATCCAAAGGATTAATAGT--- 4209
QY 181 PheGlyPheClyAsp---PheValValProValThrGluGluGlyMetAlaPheSer 199
Db 4210 TTTAATTATGGTGAATGTTGGTATTCCAGCTAGTGCAGTCAGCTCTAGCATATCAA 4269
QY 200 ThrValAsnAsn---GlyValAsnLeuGluThrAsnGluAsnPheAsp---LysIlePro 217
Db 4270 GTTCAACAGAAATATGAGTGTATTGCTGAAACAGGAGAAATATGTAATAAAGTGCCA 4329
QY 218 AlaIleAsnLeuTyr-----GlyGlyAspGlyAsnGlyGlyGlyAsnCys 232
Db 4330 CCATTATTGATTATGTTGTCATGTAGCGCGGAGGTTTAGAATAGAGTGGCTTTGTT 4389
QY 233 PheProPro-----LeuThrValPro 239
Db 4390 AAAAAATCCATGGAATGCGGAATATPATAGTGTGGAAGGATTCGCTATTAAATTCCT 4449
QY 240 LeuThrIleAsnGlnSerGlnGluLysArg-----AspValGlyLeuSerGlyGlyGlu 257
Db 4450 GTTCAAAAATCAGTTGCTCAAAATAACAATTTAGCGCTAAAGATCTGGCAGGGCTAGT 4509
QY 258 AspValGlyAspAsnValTyrProVal-----ArgMetThrValPheIleAsn 273
Db 4510 TGTGTAGTGGTATATTCACACATACCAAGTTATATCCAAACAAACAAATCACTATGGC 4569
QY 274 GluMetProIleGluValValSerGlyLeuPhe-----AsnValIleAla 288
Db 4570 GAGATTACAAATGAATTTGTAAGTGGAGCTTCAGTTACTGTGTGTGTGGAGCTTATGGA 4629
QY 289 AlaPheGlyAsnAspAlaValIleAsnSerPheGlyGlnProIle-----Leu 305
Db 4630 GCAGTACAAACGTAAGTGTTCATATTAGTAATAAAGCACTCTGTGTAAGAGGTACGTGG 4689
QY 306 ThrAspGluPheGlyVal 311
Db 4690 GCCTCTGAGTTAGGAGTA 4707

```

RESULT 2

```

US-09-402-929-1
; Sequence 1, Application US/09402929
; Patent No. 6410825
; GENERAL INFORMATION:
; APPLICANT: Temple University - Of The Commonwealth System of Higher Education
; APPLICANT: Toscani, Antonio
; APPLICANT: Hatton, Kimi
; APPLICANT: Reddy, E. P.
; TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND
; NUMBER OF SEQUENCES: 7
; TITLE OF INVENTION: USES THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/402.929
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/06896
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-214 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3602 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-402-929-1
;
Alignment Scores:
Pred. No.: 0.137
Score: 104.00
Percent Similarity: 36.18%
Best Local Similarity: 23.08%
Query Match: 5.94%
DB: 4
Length: 3602
Matches: 81
Conservative: 46
Mismatches: 156
Indels: 68
Gaps: 17
;
US-09-787-737-2 (1-325) x US-09-402-929-1 (1-3602)
QY 10 SerMetPheLysSerLysProCysAsnAsnHisHisGlnHisGlu----- 26
Db 869 TCAAAACCTTCACACAAACCTTGTGGACATATGACCATTTGCAAAACCCAGAAATCAGTTT 928
QY 27 -----IleAspThrProSerTyrMetHisTyrSer-----AsnCys----- 38
Db 929 TACATTCCTCTGATGATCCCTGGGTATCATGATGTGCGCTGATGCAATTTGTGTGAA 988
QY 39 ---AsnLeuSerSerSerPheSerSerAspArgIleProAspProLysProArgTyrAsn 57
Db 989 CATGTTTCAGACATCTGCTTTATTTCACCAACCTTTGTTGATGAAGATCCT----- 1039
QY 58 ProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnPro 77
Db 1040 GATAAAGAAAAAATAAAGAGCTCGAGTGTCTTCTTATGTCAGCC----- 1087
QY 78 ProArgGluGluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAsp 97

```



178	Gln-----	SerAlaPheGlyPheGlyAspPheValVal-----	ProValValValThr-----	192
	:::  :::	:::  :::	:::  :::	
4619	AGTGATGTTACCAAGTTTTCATCTTCTTGATGCTGCTCTCTCTCTCAAGACTACTCCCA	4678		
193	-----	-----	-----	204
	:::	-----	-----	
4679	GTATAAATTAAATGAGAATTCAACATAATGAAGAGCCATCGAATGCCAGTTTAACTCAGT	4738		
205	ValAsnLeuGlnThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrGlyGly	224		
	:::	:::	:::	
4739	CTTGACTTGAACGGGAAAAGAAC-----	4762		
225	AspGlyAsnGlyGlyAsnCysPheProProLeuThr-----	237		
	:::	:::	:::	
4763	AGTCGTAATGGTGGAGACACTGAAGCTATTCTTTTAACTCCCAAAATCGTCAAGTTT	4822		
238	---ValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGly	256		
	:::	:::	:::	
4823	AGCACTCTCCCAACCATCTCTCAGAAAAGAAAAGAAATTCGATGGGTCTCAGTCTGCAGGC	4882		
257	GluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGluMetPro	276		
	:::	:::	:::	
4883	AGTGAGCTTGGCGATGGCTCACTTAGCGAAGTTGGTAATGCAGACATCAACAACACACCA	4942		
277	IleGluValVal-----	SerGlyLeuPheAsnValIlysAlaIlePheGlyAsn	292	
	:::	:::	:::	
4943	GTGAACAACACTACCATTTTCTCTCTCAGTTTTTTTAAAC-----	ACATGCTCGTGGAAAT	4996	
293	AspAlaValIleuIleAsn-----	SerPhe-----	GlyGlnProIleLeuThrAspGluPhe	309
	:::	:::	:::	
4997	GAACAACATTAATATAGAAAACCCCTCTCTCTTACATCAACCCCAATTTGTGGGCAGAAAGTT	5056		
310	GlyValThrTyrGlnProLeuGlnAsnGlyAla	320		
	:::	:::	:::	
5057	CTCATCTTACAA-----	ACTGCTCTTTTCAGAAAGAAC	5086	

```

RESULT 4
; US-09-031-962D-3
; Sequence 3, Application US/09031962D
; Patent No. 6350867
; GENERAL INFORMATION:
; APPLICANT: Thomas C. Hart
; APPLICANT: Jennifer A. Price
; TITLE OF INVENTION: Methods and Compositions for Enhancing
; TITLE OF INVENTION: Osseous Growth, Repair, and Regeneration
; FILE REFERENCE: WFO98-18
; CURRENT APPLICATION NUMBER: US/09/031,962D
; CURRENT FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-031-962D-3

Alignment Scores:
Pred. No.: 0.017 Length: 860
Score: 103.50 Matches: 46
Percent Similarity: 40.41% Conservative: 32
Best Local Similarity: 23.83% Mismatches: 59
Query Match: 5.91% Indels: 56
DB: 4 Gaps: 9

US-09-787-737-2 (1-325) x US-09-031-962D-3 (1-860)

QY 21 HisHisHisGlnHisGluLeuAsp-----ThrProSerTyrMethHisTyrSer 36
:::|||||:::
Db 193 TACCACCACCAATTCATATCTCAATGGGCTTGACGGCACCGGCGCTTACTGCGCCCAAGTCG 252
:::|||||:::

QY 37 AsnCysAsnLeuSerSerSerPhe-----
:::|||||:::
Db 253 GAATATACCTACGGAGCCTCTCTCTACGGCAATACGGGCGGTATCGGGAGCAGCGCTGCCA 312
:::|||||:::

```

QY	45	-----SerSerAspArgIleProAspProLysProArgTrp--AsnPro	58
Db	313	GCCAGGACCCAGTGCCTGAAGGAGGACC CGAAGCAGAGTGCGCATGGTGAATGGG	372
QY	59	LysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProPro	78
Db	373	AAGCCCAGAAGTCGGAAGCCGCTACAATTACTCCAGTACCAGTGGCGGCCCTG	432
QY	79	ArgGluGluIleGlnArgIleAArgIle-----ArgLeuGlnGluTyrgly	93
Db	433	CAGCGCGCTTC CAGAAGGCCCACTGCTGGCGGTGCCGAGCGCGGAGCTGGCCGCG	492
QY	94	GlnIleGly-----AspAlaAsnValPheTyrtipPheGlnAsnAArgLysSerArgAla	111
Db	493	CAGCTGGCGCTCACGCACACACAGGTGAATACTGGTTCCAGAACCCGCGTTCCAAAGTTC	552
QY	112	LysHisLysLeuArgValHisLysSerProLysMetSerLysLysAspLysThrVal	131
Db	553	AAG-----AAACTCTCAGAAACAGGTGCGCTGGAGCAGATCCCCAATACAGTG	603
QY	132	IleProSerThrAspAlaAspHisCysPheglyPheValAsnGlnGluThrGlyLeuTyrr	151
Db	604	ATTCCA-----TGCG	612
QY	152	ProValGlnAsn-AsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHi	171
Db	613	CCTGCACTCACCACC-----ATCACCGCGCCTCTGGAGACACCTCTTCCCACTCCA	663
QY	171	sAsnAspPro-----SerAlaAlaGlnSerAla	180
Db	664	CTCCGGGCGCTTGGCCGAGTCAGTGGCCCGCGCGT	700

```

RESULT 5
US-09-489-039A-619/c
; Sequence 619, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 619
; LENGTH: 1404
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-619

Alignment Scores:
Pred. NO.: 0.0476 Length: 1404
Score: 102.50 Matches: 52
Percent Similarity: 39.64% Conservative: 36
Best Local Similarity: 23.42% Mismatches: 69
Query Match: 5.85% Indels: 65
DB: 4 Gaps: 12

US-09-787-737-2 (1-325) x US-09-489-039A-619 (1-1404)

QY 107 ArgLysSerArgAlaIysHisLys-----LeuArgValHisHis 119
|||:||||:||||:||||:||||:
Db 801 CGCGGGAATGTTGGCGGCGACACGCGATCCGACCGCGGTTGATCTCGGCTATCACCTC 742
||||:||||:||||:||||:||||:
QY 120 LysSerPro-----LysMetSerLysLysAspLysThrValIlePro 133
::: |||: |||: |||: |||: |||:
Db 741 CGGCAGCAATCTTTTACCTTTGTGCAGGATCTGGCGCACGAAGATAAA---GTGTTTCAG 685
|||: |||: |||: |||: |||:
QY 134 SerThrAspAlaAspHisCysPheValAsn-----GlnGluThrGlyLeuTyr 151
::: |||: |||: |||: |||: |||:

```

```

Db      684  GCGGCGACCTGGATCACCACCTCTTTTCTGTCAGCGCATTCGACGAGTGT---TAT 628
Qy      152  ProValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHis 171
Db      627  CCCCTTTTGCATATACCCGCGACGTTACACAGCCCGACGCGCTTT----- 583
Qy      172  AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValValProValVal 191
Db      582  ---GATCGGCTGTGGCGCGAGATCGCTTCGCTCACCATCCCGACCGGTGTGTAAGTT 526
Qy      192  ThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGlu 211
Db      525  CAGCAGCCGCGGTGGGACACAGTCGCTCCATATCGCGACGATCTCCAG----- 475
Qy      212  AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGly---Gly 230
Db      474  -----GGCGATGGGAATGTGCGACAGGC 451
Qy      231  AsnCysPheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys----- 248
Db      450  GTTGGGAAGCCGCCCAAGGCCATTGTCTCTCGCGCATCAAGCCCATACTTGAGTGAGAT 391
Qy      249  -----ArgAspValGlyLeuSerGly-----GlyGluAspVal 259
Db      390  ACCTTCATCGCTGATCGCGGCGTCAAGCAGCGCGCGGCGAACTGCGAGCAGCAGAAATC 331
Qy      260  Gly-----AspAsnValTyrProValArgMetThrValPheIleAsnGluMetPro--- 276
Db      330  GGCATCGCGCAGCGCGCTTTCGCGTCAAGGTGGCCACCATTCATAGTCAGCGCGGC 271
Qy      277  -----IleGluValValSerGlyLeuPhe 284
Db      270  TTTGGCGATCATTCGCGCGCGCGAGTCGCGCGATAATCTCGACTTCTCTTTCGCTCTTC 211
Qy      285  AsnVal 286
Db      210  AATGTC 205

RESULT 6
US-09-322-478-17
; Sequence 17, Application US/09322478
; Patent No. 6331662
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; APPLICANT: Voycas, Daniel F.
; TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
; FILE REFERENCE: P-1065 ISURF Plant Retroelement
; CURRENT APPLICATION NUMBER: US/09/322,478
; CURRENT FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087125
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 12286
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plant
US-09-322-478-17

Alignment Scores:
Pred. No.: 1.78 Length: 12286
Score: 101.50 Matches: 76
Percent Similarity: 33.03% Conservative: 33
Best Local Similarity: 23.03% Mismatches: 96
Query Match: 5.80% Indels: 125
DB: 4 Gaps: 16

US-09-787-737-2 (1-325) x US-09-322-478-17 (1-12286)
Qy      20  AsnHisHisGlnHisGluIleAspThrProSerTyrMethHisTyrSer----- 36

```

```

Db      2680 AACCAACATACCAACCCACCATACCAAGCACCCTACCAAGTCCGAATCAGCAAGAAAGC 2739
Qy      37  -----AsnCysAsnLeuSerSerPheSerSerPheSerSerPheSerSerPheSerSerPhe 51
Db      2740 CCACCAAAATAGAGAACTCTGCTGCAATTCATCAGAGAGACAAAGATCATCATCAAAAGA 2799
Qy      52  -----Prolys 53
Db      2800 GCACGGATGCAGCCATTCCGAATCTAGAGTTCAATGGGCCCACTGGCCGATGACAAAG 2859
Qy      54  ProArg-----TyrAsnProLysProGluGlnIleArg 64
Db      2860 CCAGACGGCCCATAGAACTTTCGGTGTAAATGAGAGAGAACCC---GAGGAGG 2916
Qy      65  Ile-----LeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGlu 80
Db      2917 ATAAACGAGTACTGACTAGAGGC-----AGAGAA 2946
Qy      81  GluIleGlnArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100
Db      2947 GAG---CGCAGGAGGAGGCTAAGGTTGAAGGAG-----AAGCTGCCAGAGAGAGAA 2997
Qy      101  PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgVal-HisHisly 120
Db      2998 GGACAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3057
Qy      120  sSerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCy 140
Db      3058 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3117
Qy      140  sPheGlyPheValAsnGlnGluThrGlyLeu-TyrProValGlnAsnAsnGlu----- 157
Db      3118 CTATGCCACCCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3177
Qy      158  -----LeuValValThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerA 176
Db      3178 AAGGGTTAGAAATCACTATGCCATTCGGG-----GAAG 3210
Qy      176  laAlaGlnSerAlaPheGlyPheGlyAspPheValProValValThrGluGluGlyM 196
Db      3211 CCTTACAGCAGATGCCCTCTACTCCAAATTTATGAAAGACATCTCCACCAAGAGAGGCA 3270
Qy      196  etAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysI 216
Db      3271 AGTATATTGAC-----AACGAGAATTTGTGTA- 3299
Qy      216  leProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGlyAsnCys----- 232
Db      3300 -----GGAGGCAATTGCAGTCGATAA 3321
Qy      233  -----PhePro-----ProLeuThrValProLeuThrI 242
Db      3322 TACAAAGGATTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3381
Qy      242  leAsnGlnSerGlnGluLysArg-----AspValGlyLeuSerGlyGlyGluAspV 259
Db      3382 TTGGGAAGGAGAGCGGTAAACAAGGCCCTCATTTGATCTAGAGAGCAAGT----- 3428
Qy      259  aGlyAspAsnValTyrProValArgMetThrValPheIleAsnGluMetProIleGluV 279
Db      3429 -----ATCAATCTGATGCCCTTGTCAA 3450
Qy      279  alValSerGlyLeuPheAsnValLys 287
Db      3451 TGTGCAAAAGAAATGGGAATTTGAAG 3476

RESULT 7
US-08-426-236-3
; Sequence 3, Application US/08426236
; Patent No. 5629188
; GENERAL INFORMATION:
; APPLICANT: Shiba, Kiyotaka

```

QY	212	AsnPheAspLysIle-----ProAlaIleAsnLeuTyr-----	222
		:::     :::	
Db	1822	TCATTGCGAATAGCTACGTCCGATTCCTCATTAAGAACAATCATACTGGTACACATATCTTA	1881
		:::     :::	
QY	223	-----GlyGlyAspGlyAsnGlnGlyGlyGlyAsnCysPhePro	234
		:::     :::	
Db	1882	AACTTTGCTCTGAAGAAACTTTGGGTATGATGTCATCAAAAAGGCTGTGTCAATGAAGAATTGAAA	1941
		:::     :::	
QY	235	ProLeuThrValProLeuThrIleAsnGlnSerGlnGlnLysArgAspValGlyLeuSer	254
		:::     :::	
Db	1942	CAGAAAAAATTGAGATTCGATTCTCTCATAAAAAGGCTGTGTCAATGAAGAATTGAAA	2001
		:::     :::	
QY	255	GlyGlyGluAspValGlyAspIleValTyrProValArgMetThrValPheIleAsnGlu	274
		:::     :::	
Db	2002	AAAGTTGAAGATATCTGTAATGAGCAAATTAAGAAAAAATCAAGTCTTTTACAGGAA	2061
		:::     :::	
QY	275	MetProIleGluValValSerGlyLeuPheAsnValLysAlaIlePheGlyAsn-----	292
		:::     :::	
Db	2062	ATTCCCAITGGACTTGGCCAAATCCAATGATGGGTTCGTCTGCTTTGGTGAGACTTAC	2121
		:::     :::	
QY	293	---AspAlaValLeuIleAsnSerPheGlyGlnProfile	304
		:::     :::	
Db	2122	CCAGATCCAGTTCGTGTAGTTTCTGTGGTGAAGCCAAATC	2160
		:::     :::	
<b>RESULT 8</b>			
US-08-781-891-207/c			
; Sequence 207, Application US/08781891			
; Patent No. 6090620			
; GENERAL INFORMATION:			
; APPLICANT: Fu, Ying-Hui			
; APPLICANT: Yu, Chang-Shen			
; APPLICANT: Oshima, Junko			
; APPLICANT: Mulligan, John T.			
; APPLICANT: Schellenberg, Gerald D.			
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO			
; TITLE OF INVENTION: WERNER'S SYNDROME			
; NUMBER OF SEQUENCES: 209			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: SEED AND BERRY LLP			
; STREET: 6300 Columbia Center, 701 Fifth Avenue			
; CITY: Seattle			
; STATE: Washington			
; COUNTRY: USA			
; ZIP: 98104-7092			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/781,891			
; FILING DATE: 27-DEC-1996			
; CLASSIFICATION: 800			
; ATTORNEY/AGENT INFORMATION:			
; NAME: No. 6090620tenburg Ph.D., Carol			
; REGISTRATION NUMBER: 39,317			
; REFERENCE/DOCKET NUMBER: 240052.419			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (206) 622-4900			
; TELEFAX: (206) 682-6031			
; INFORMATION FOR SEQ ID NO: 207:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 29604 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-08-781-891-207			
<b>Alignment Scores:</b>			
Pred. No.:	26.5	Length:	29604
Score:	96.50	Matches:	71
Percent Similarity:	38.43%	Conservative:	32

Best Local Similarity: 26.49% Mismatches: 90  
 Query Match: 5.51% Indels: 75  
 DB: 3 Gaps: 17

US-09-787-737-2 (1-325) x US-08-781-891-207 (1-29604)

```

QY 52 ProLysProArgTTPAsnPro---LysProGluGlnIleArg----- 64
Db 22978 CCTAAACACAGAAAGACCCACAAAGATAGAGAACTTCAGACCAATTTCTCTTATGAAT 22919
QY 65 IleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluIleGlnArg 84
Db 22918 ATCATGCAAAATCTCAATAAAATTCCTGCTAACCGAATCCAGAAACACATTAAGCA 22859
QY 85 IleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrpPhe 104
Db 22858 ATC---ATCCATCTGACCAAGTAGGTTTATT-----CCAGGGATGCGAGGATGTTT 22808
QY 105 GlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLysMet 124
Db 22807 AATATACGAAATCC-----ATCAATGTAATCCATTATATAAACAATC 22763
QY 125 SerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPheVal 144
Db 22762 AAAGACAAAACACATGATCATC---TCGTTAGATGCAAGAAACCATTTGCAAGATC 22706
QY 85 IleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrpPhe 104
Db 22858 ATC---ATCCATCTGACCAAGTAGGTTTATT-----CCAGGGATGCGAGGATGTTT 22808
QY 105 GlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLysMet 124
Db 22807 AATATACGAAATCC-----ATCAATGTAATCCATTATATAAACAATC 22763
QY 125 SerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPheVal 144
Db 22762 AAAGACAAAACACATGATCATC---TCGTTAGATGCAAGAAACCATTTGCAAGATC 22706
QY 145 AsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValThrGluProAla 164
Db 22705 CAA-----CACCATTCATG-----ATAAAGTGTGGAAAGATCA 22670
QY 165 GlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPheGly 184
Db 22669 GGAATTCAGGCCCATACCTAAACATGATAAAGCAATCTTACAGCAA----- 22622
QY 185 AspPheValProValThrGluGluGlyMetAlaPheSerThrValAsnAsnGly 204
Db 22621 -----CCAGTA----- 22616
QY 205 ValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrGlyGly 224
Db 22615 GCCAATCAAGGCCCATACCTAAACATGATAAAGCAATCTTACAGCAA----- 22622
QY 225 AspGlyAsnGlyGlyGlyAsnCysPheProLeuThrValProLeu----- 240
Db 22561 ACTAGACAGGC-----TGC-----CCACTTCTCCTACCTTTTCAACATAGTA 22517
QY 241 -----ThrIleAsnGlnSerGlnGluLysArgAspValGlyLeuSer 254
Db 22516 CTTGAAGTATTAGCCAGAGCAATTCGACAGAAAGGAGATCAAGGGGATACAAATT--- 22460
QY 255 GlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsnGlu 274
Db 22459 GGAAGAGAGGAAGTCAAAATATCATCTTTTCAGATGATATAGTATATATAGTGAC 22400
QY 275 MetProIleGluValValSerGlyLeuPheAsnValLysAlaIlePheGlyAsnAspAla 294
Db 22399 ---CCTAAATTTACAGAGAACTCTCTAAC----- 22370
QY 295 ValLeuIleAsnSerPheGlyGln 302
Db 22369 ---CTGATAACAGCTTCGGTGAA 22349

```

## RESULT 9

US-09-618-166-207/c

Sequence 207, Application US/09618166

Patent No. 6583112

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

Yu, Chang-En

Oshima, Junko

Mulligan, John T.

Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

## WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/618,166

FILING DATE: 17-Jul-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: McWaters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 240052.419C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 207:

SEQUENCE CHARACTERISTICS:

LENGTH: 29604 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 207:

US-09-618-166-207

## Alignment Scores:

Pred. No.:	Score:	Length:
26.5	36.50	29604
Percent Similarity:	38.43%	Matches:
Best Local Similarity:	26.49%	Conservative:
Query Match:	5.51%	Mismatches:
DB:	4	Indels:
		Gaps:
		17

US-09-787-737-2 (1-325) x US-09-618-166-207 (1-29604)

```

QY 52 ProLysProArgTTPAsnPro---LysProGluGlnIleArg----- 64
Db 22978 CCTAAACACAGAAAGACCCACAAAGATAGAGAACTTCAGACCAATTTCTCTTATGAAT 22919
QY 65 IleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluIleGlnArg 84
Db 22918 ATCATGCAAAATCTCAATAAAATTCCTGCTAACCGAATCCAGAAACACATTAAGCA 22859
QY 85 IleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrpPhe 104
Db 22858 ATC---ATCCATCTGACCAAGTAGGTTTATT-----CCAGGGATGCGAGGATGTTT 22808
QY 105 GlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLysMet 124
Db 22807 AATATACGAAATCC-----ATCAATGTAATCCATTATATAAACAATC 22763
QY 125 SerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPheVal 144
Db 22762 AAAGACAAAACACATGATCATC---TCGTTAGATGCAAGAAACCATTTGCAAGATC 22706
QY 145 AsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValThrGluProAla 164
Db 22705 CAA-----CACCATTCATG-----ATAAAGTGTGGAAAGATCA 22670
QY 165 GlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPheGly 184
Db 22669 GGAATTCAGGCCCATACCTAAACATGATAAAGCAATCTTACAGCAA----- 22622
QY 185 AspPheValProValThrGluGluGlyMetAlaPheSerThrValAsnAsnGly 204

```







1594	Db	CTTCTATCA-----				-----	CCAAAACCTCCTCCTCTGTTCTTTCTCCT	1556
55	Qy	-----	-----	-----	-----	-----	-ArgTspAsnProLysPro	60
1555	Db	CCTTGTCTCTCCATTTCTCGGCTCTTAACAACATAC				-----	TCTGTTATCGAATCCAAACCA	1496
61	Qy	GlucGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnPro-----				-----		77
1495	Db	AAGGAAT-----				-----	TCCATGGCTTTGAATCAAAAGTGAACCCCTTTTGGCTT	1448
78	Qy	-----	-----	-----	-----	-----	-ProArgGluGluIleGlnArg	84
1447	Db	CTCTTAGTAGCTGCTTTTCGCGATGTGGTGGCTTCGCAACCAACCACTGAAGATCAACAT				-----		1388
85	Qy	IleArgIleArgLeuGlnGluTyr-----				-----	-GlyGlnIleGlyAspAlaAsnValPhe	101
1387	Db	CTTCGTAAACATACCCCAATCTTTTAGTATCAAAAGCAATCTCTTACCTTCCACGCGAACAA				-----		1322
102	Qy	TyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgVal--HisHisLys				-----		120
1327	Db	AATGGCAATTCATCATAGAAGTAAAAATCAGAATCTAAACCAATCTGGATCATCACATAT				-----		1261
121	Qy	SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys				-----		140
1267	Db	AGTCCTTTCTTCTCTCCCAATTTTTTCCATATCCCCCACTCCGATCCGATCGGATTCGTGT				-----		1201

RESULT 14  
US-09-031-962D-1

```

US-09-031-962D-1
; Sequence 1, Application US/09031962D
; Patent No. 6350867
; GENERAL INFORMATION:
; APPLICANT: Thomas C. Hart
; APPLICANT: Jennifer A. Price
; TITLE OF INVENTION: Methods and Compositions for Enhancing
; TITLE OF INVENTION: Ossous Growth, Repair, and Regeneration
; FILE REFERENCE: WPU98-18
; CURRENT APPLICATION NUMBER: US/09/031,962D
; CURRENT FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-031-962D-1

```

Alignment Scores:			
Pred. No.:	0.285	Length:	864
Score:	93.00	Matches:	32
Percent Similarity:	40.62%	Conservative:	36
Best Local Similarity:	23.96%	Mismatches:	61
Query Match:	5.31%	Indels:	54
Da:	4	Gaps:	9

US 00 707-727-2 (1-325) x US-09-031-962D-1 (1-864)

Qy	21	HisHisGlnHisGluLeuAsp-----ThrProSerTyrMethHisTyrSer	36
		:::     :::     :::     :::	
Db	193	TACCACCACCAATTCAATCTCAATGGGCTGCAGGACGGGGCTTACTCGCCCAAGTCG	252
Qy	37	AsnCysAsnLeuSerSerSerPhe-----	44
		:::     :::     :::     :::	
Db	253	GAATATACCTACGGAGCGCTCTACCGGCAATACGGGGCGTATCGGAGACGCGCTGCCA	312
Qy	45	-----SerSerAspArgileProLysProArgTrp-----AsnPro	58
		:::     :::     :::     :::	
Db	313	GCCACGAGCCCGTCTCGTGAAGGAGAGCGCGGAGACAGAGTGGCGATGTGAATGGG	372
Qy	59	LysProGluGlnIleArgileLeuGluSerIlePheAsnSerGlyThrIleAsnProPro	78
		:::     :::     :::     :::     :::     :::	
Db	373	AAGCCCAAGAAAGCTCCGAAGCCGGGTCAATCTACTCCAGCTACCAAGTGGCGGCCCTG	432

Db	3188	CTATGCCATTCCGGG-----				GAAGCTTTACAGCAGATGC	3220
Qy	181	heGlypheGlyAspPheValValProValValThrGluGluGlyMetalPheSerThrV					201
Db	3221	CCCTCTACTCCAAATTTATGAAAGACATCTCTCCAGAGAGGGGAGTATTATGAC----					3276
Qy	201	alaAsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnL					221
Db	3277	-----AACGAGAAATATTGTGGTA-----					3294
Qy	221	eutYrGlyGlyAspGlyAsnGlyGlyGlyAsnCys-----				PheP	234
Db	3295	-----GGAGGCAATTGCAGTGCAGTAATAACAAAGGAAGCTAC					3331
Qy	234	ro-----ProLeuThrValProLeuThrIleAsnGlnSerGlnG					247
Db	3332	CCAAAGAGTTTTAAAGACCCCGAAGTGTACTCATCCCGTCGCCAAT-----					3378
Qy	247	luLysArgAspValGlyLeuSerGlyGlyGluAspValGlyAspAsnValtyProVala					267
Db	3379	-----GGGAGGAAGCCGTAAACAAGGCCCTCATTTGATCTAA					3415
Qy	267	rgMetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal					287
Db	3416	GAGCAAGT-----ATCAATCTGATGCCCTTGTCATATGTCACAAAGAAATTTGGAATTTGA					3469
Qy	287	ys					287
Db	3470	AG					3471

RESULT 13

US-09-889-463A-35/C  
Sequence 35 Application US/09889463A

```

sequence 35, application 08/790000000.
Patent No. 6680185
GENERAL INFORMATION:
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Falco, Saverio C.
• APPLICANT: Falco, Saverio C.
APPLICANT: Kinney, Anthony J.
APPLICANT: Miao, Guo-Rua
TITLE OF INVENTION: Plant Polyphenol Oxidase
FILE REFERENCE: BB1330
PUBLICATION NUMBER: US/09/889,463A

```

; CURRENT APPLICATION NUMBER: US/03/800,000  
 ; CURRENT FILING DATE: 2001-07-16  
 ; PRIOR APPLICATION NUMBER: 60/119,590

; PRIOR APPLICATION NUMBER: 00/115,035  
 ;  
 ; PRIOR FILING DATE: 1999-02-10  
 ;  
 ; NUMBER OF SEQ ID NOS: 46

```

; NUMBER OF SEQ ID NOS. 10
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 35

```

```

; SEX: 10
; LENGTH: 2260
; TYPE: DNA

```

ORGANISM: Glycine max  
US-09-889-463A-35

**Alignment Scores:**

Pred. No.:	1.1
Score:	93.50

Percent Similarity:	36.67%
Best Local Similarity:	21.11%
	5.34%

Query Match: 5.34%

DB: 4

US-09-787-737-2 (1-325) x US-09-889-463

```
QY      2 SerSerSerAsnLysAsnTrpPro
      : ::::: |:::|
      : ::::: |:::|
```

Db 1711 GCTGCACATTCAAAAACCTCCCA

Qy 21 ---HisHisHisGlnHisGluIle

Db 1654 CTTTCATCATCATCAATATGAACAT

QY 40 LeuSerSerSerPheSerSerAsp

```

QY 79 ArgGluGluIleGlnArgIleArgIle-----ArgLeuGlnGluTyrGly 93
Db 433 CAGCGCGCTCCAGAGGCCAGTACCTGGCGTCCCGCGCGCGAGCTGGCGCG 492
QY 94 GlnIleGly-----AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAla 111
Db 493 CAGTGGGCTCAGCGCAGACACAGGTGAATCTGGTTCCAGAACCCGCGTCCAAAGTTC 552
QY 112 LysHisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrVal 131
Db 553 AAG---AAACTCTACAGAAGCGGAGGTGGCGTGGAGCACACTCCCAAT- 600
QY 132 IleProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyr 151
Db 601 -----AACAGTGATTCCTAGGCGTGC----- 621
QY 152 ProValGlnAsnAsnGluLeuValThrGluProAlaGlyPheLeuPheProValHis 171
Db 622 -----AACTCACC-ACCATCACCGCGCTCTGGGACACCTTTCCCACTCCAC 668
QY 172 AsnAspPro-----SerAlaAlaGlnSerAla 180
Db 669 TCGGCGCCCTGCGCGCAGTCAGCTGCGCGCGCGCT 704

RESULT 15
US-09-252-991A-14226
; Sequence 14226, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14226
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14226

Alignment Scores:
Pred. No.: 0.369 Length: 936
Score: 92.50 Matches: 40
Percent Similarity: 39.88% Conservative: 27
Best Local Similarity: 23.81% Mismatches: 62
Query Match: 5.28% Indels: 39
DB: 4 Gaps: 9

US-09-787-737-2 (1-325) x US-09-252-991A-14226 (1-936)
QY 112 LysHisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrVal 131
Db 247 CGGCATGCTGTGGACAACACACCGCGCTGGCGCACCGGCAACCGGCAACGACAGTGT 306
QY 132 Ile-----ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGlu 147
Db 307 GTATCCCATCTACGGCTGTGGCGGCGCACAGACAGCC-----GGTCCGCGGAC 357
QY 148 ThrGlyLeuTyrProValGlnAsnGluLeuValThrGluProAlaGlyPheLeu 167
Db 358 CTACCGGACACCGTCAACGTGACCATCAGTTGGTGGATGCCGATGTCCAGCGCGCTCTA 417
QY 168 Phe-----ProValHisAsnAspProSerAla 176
Db 418 TTCTGTGCTTCTCCTGTGCTGCGGAACTGTCGCGCATCGCATCCATCGCG 477
QY 177 AlaGlnSerAlaPheGlyPheGlyAspPheValProValThrGluGluGlyMet 196

```

```

Db 478 ATCAGCGCGCGC-----TTCAGGTCCAGGCGGTGGTCCCAACGCGTGC 522
QY 197 AlaPheSerThrVal-----AsnAsnGlyValAsnLeuGluThrAsnGluAsnPheAsp 214
Db 523 GCCTTCGGCACCGCATCAGCAACACGCTATGACCTCGGCACC--CTGAGCTTCGGC 579
QY 215 LysIle-----ProAlaIleAsnLeuTyrGlyGlyAspGlyAsnGlyGly--- 229
Db 580 ACCCTCGGCAACCTGGCCAGTCCGGTTAAGTCCCGAGCAGTCCGGGGCGCGGTCCATC 639
QY 230 ---GlyAsnCysPheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys 248
Db 640 GTCTGACCTGCACTCCGGGAATGACCGTCTCGGTGCGCCCTC----- 681
QY 249 ArgAspValGlyLeuSerGlyGly 256
Db 682 ---GACTACGGGTCAATGGCGGC 702

```

Search completed: September 10, 2004, 12:00:29  
Job time : 4581 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 10, 2004, 10:47:43 ; Search time 564 Seconds  
(without alignments)  
2899.370 Million cell updates/sec

Title: US-09-787-737-2  
Perfect score: 1751  
Sequence: 1 MSSNNKWSMPKSKPCNNN.....TDFGVYQPLQNGAIYVLI 325

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool\_h/US09787737/runat\_01092004.112458.23693/app\_query.fasta\_1.519  
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnnp -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=-1 -MATRIX=blowsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09787737 @CEN 1 1 520 @runat\_01092004.112458.23693  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADS=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DSLOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

1	623	35.6	1502	13	US-10-225-066A-329	Sequence 329, App
2	623	35.6	1502	16	US-10-374-780A-2863	Sequence 2663, Ap
3	598	34.2	2164	13	US-10-424-599-129204	Sequence 129204,
4	427	24.4	1961	13	US-10-425-114-27293	Sequence 27293, A
5	405	23.1	2066	17	US-10-437-963-61875	Sequence 61875, A
6	394.5	22.5	807	10	US-09-934-455-243	Sequence 243, App
7	394.5	22.5	807	13	US-10-225-066A-165	Sequence 165, App
8	366	20.9	1524	17	US-10-374-780A-2665	Sequence 2665, Ap
9	354.5	20.2	1215	13	US-10-437-963-59649	Sequence 59649, A
10	354.5	20.2	1215	13	US-10-425-114-12809	Sequence 12809, A
11	342	19.5	994	10	US-09-934-455-489	Sequence 97798, A
12	338	19.3	837	16	US-10-260-238-1567	Sequence 489, App
13	338	19.3	837	17	US-10-437-963-43264	Sequence 1567, Ap
14	329.5	18.8	921	17	US-10-437-963-59651	Sequence 43264, A
15	328.5	18.8	1367	17	US-10-437-963-59651	Sequence 59651, A
16	304.5	17.4	789	16	US-10-260-238-5518	Sequence 5518, Ap
17	291.5	16.6	1279	13	US-10-437-963-85739	Sequence 85739, A
18	212.5	12.1	816	13	US-10-425-114-441	Sequence 331, App
19	212.5	12.1	816	16	US-10-225-066A-331	Sequence 441, App
20	212.5	12.1	816	13	US-10-374-780A-2667	Sequence 331, App
21	199.5	11.4	1227	13	US-10-412-699B-919	Sequence 2667, Ap
22	199.5	11.4	1227	13	US-10-225-066A-323	Sequence 919, App
23	199.5	11.4	1227	16	US-10-374-780A-301	Sequence 323, App
24	193.5	11.1	2212	17	US-10-744-572-77	Sequence 301, App
25	193.5	11.1	2212	17	US-10-437-963-19162	Sequence 77, Appl
26	188	10.7	1154	13	US-10-425-114-15010	Sequence 19162, A
27	188	10.7	1289	13	US-10-424-599-113771	Sequence 15010, A
28	188	10.7	1289	13	US-10-744-572-76	Sequence 113771,
29	187.5	10.7	1338	17	US-10-744-572-3	Sequence 76, Appl
30	187.5	10.7	1338	17	US-10-744-572-3	Sequence 3, Appl
31	187	10.7	741	17	US-10-424-599-116468	Sequence 116468,
32	187	10.7	900	13	US-10-424-599-113772	Sequence 113772,
33	186.5	10.6	981	13	US-10-424-599-113772	Sequence 28807, A
34	185	10.6	1186	13	US-10-425-114-28807	Sequence 23553, A
35	184.5	10.5	731	13	US-10-437-963-23553	Sequence 10583, A
36	183.5	10.5	731	13	US-10-424-599-10583	Sequence 1680, Ap
37	183.5	10.5	731	13	US-10-412-699B-1680	Sequence 1669, Ap
38	183.5	10.5	731	16	US-10-374-780A-1680	Sequence 35592, A
39	183	10.5	2014	13	US-10-425-114-35592	Sequence 24906, A
40	181	10.3	998	13	US-10-424-599-129203	Sequence 21484, A
41	179	10.1	423	17	US-10-424-599-129203	Sequence 107605,
42	177	10.1	422	13	US-10-767-701-21484	Sequence 129203,
43	176	10.1	688	13	US-10-424-599-129203	Sequence 17, Appl
44	175.5	10.0	844	17	US-10-744-572-17	Sequence 291, App
45	175.5	10.0	870	17	US-10-437-963-291	

ALIGNMENTS

RESULT 1  
US-10-225-066A-329  
; Sequence 329, Application US/10225066A  
; Publication No. US20030226173A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendel Biotechnology, Inc.  
; APPLICANT: RATCLIFFE, Oliver  
; APPLICANT: RIECHMANN, Jose Luis  
; APPLICANT: ADAM, Luc J  
; APPLICANT: DUBELL, Arnold T  
; APPLICANT: HEARD, Jacqueline E  
; APPLICANT: PILGRIM, Marsha L  
; APPLICANT: JIANG, Cai-Zhong  
; APPLICANT: REUBER, T. Lynne  
; APPLICANT: CREELMAN, Robert A  
; APPLICANT: PINEDA, Omaira  
; APPLICANT: YU, Guo-Liang  
; APPLICANT: BROUN, Pierre E  
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants  
; FILE REFERENCE: MB10036-2 US  
; CURRENT APPLICATION NUMBER: 2002-08-09  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 09/837,444

RESULT 2  
US-10-374-780A-2663  
Sequence 2663, Application US/10374780A  
Publication No. US20040019927A1  
GENERAL INFORMATION:  
APPLICANT: Sherman, Bradley K  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Heard, Jacqueline E  
APPLICANT: Haake, Volker  
APPLICANT: Creelman, Robert A  
APPLICANT: Ratcliffe, Oliver  
APPLICANT: Adam, Luc J  
APPLICANT: Reuber, T. Lynne  
APPLICANT: Keddle, James  
APPLICANT: Broun, Pierre E  
APPLICANT: Pilgrim, Marsha L  
APPLICANT: Dubell III, Arnold T  
APPLICANT: Pineda, Omalra  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
FILE REFERENCE: MBI-0047 CIP  
CURRENT APPLICATION NUMBER: US/10/374,780A  
CURRENT FILING DATE: 2003-02-25  
PRIOR APPLICATION NUMBER: 09/837,944  
PRIOR FILING DATE: 2001-04-18  
PRIOR APPLICATION NUMBER: 60/310,847  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 09/934,455  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/336,049  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/338,692  
PRIOR FILING DATE: 2001-12-11  
PRIOR APPLICATION NUMBER: 10/171,468



:	PRIOR FILING DATE:	2002-06-14				
:	PRIOR APPLICATION NUMBER:	10/225,066				
:	PRIOR FILING DATE:	2002-08-09				
:	PRIOR APPLICATION NUMBER:	10/225,067				
:	PRIOR FILING DATE:	2002-08-09				
:	PRIOR APPLICATION NUMBER:	10/225,068				
:	PRIOR FILING DATE:	2002-08-09				
:	NUMBER OF SEQ ID NOS:	2906				
:	SOFTWARE:	PatentIn version 3.2				
:	SEQ ID NO	2663				
:	LENGTH:	1502				
:	TYPE:	DNA				
:	ORGANISM:	Arabidopsis thaliana				
:	FEATURE:					
:	OTHER INFORMATION:	G1584				
:	US-10-374-780A-2663					
Alignment Scores:						
Pred. No.: 5,27e-69 Length: 1502						
Score: 623.00 Matches: 164						
Percent Similarity: 49.64% Conservative: 40						
Best Local Similarity: 39.90% Mismatches: 83						
Query Match: 35.58% Indels: 124						
DB: 16 Gaps: 18						
US-09-787-737-2 (1-325) x US-10-374-780A-2663 (1-1502)						
QY	1	MetSerSerSerAsnLysAsnTrpProSerMetPheLysSerLysProCysAsnAsnAsn	20			
Dd	160	ATGGCTCTTCGATAGACACTGGCCACAGTGTCAAGTCCAACCT-----CATCCC	213			
Q*	21	HISHisGlnHISglurleAspThrProSerTyrMetHisTy-SerAsnCysAsnLeu	40			
Dd	214	CATCAATGCCAACATGACATCAACTCTCT-----CTCTGCCTTCGTCTTCTCACCGA	267			
QY	41	SerSerSerPheSerSer-----AspArgIleProAspProLysProArgTrp	56			
Dd	268	TCTTCTCTCTTTCTCTTCAGGATGTCAAGTGGAGAGAGTCCAGAGCCAAACAAGATGG	327			
QY	57	AsnProLysProGluGlnIlearglleuGluSerIlePheAsnSerGlyThrIleAsn	76			
Dd	328	AATCCAAAGCCAGACAGCATTCGGATATTGAAGCAATCTTTAACTCCGGGATGTGAAT	387			
QY	77	ProProArgGluGluIleGlnArgIlearglleuGluGlnGluTyrglyGlnIleGly	96			
Dd	388	CCTCCAAAGAGAGGAG-----ATCAGGCTTCAGAATAACGCCCAATGTCGGT	432			
QY	97	AspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg	116			
Dd	433	GATGCTAACGCTTCTACTTGTTCCAAAACCGTAAGTCCCGTAGTAAACACAAACTCCG	492			
QY	117	ValHisHis-----	119			
Dd	493	CTCTCCACACACACTCCAAACACTCTCTCCCTCAAACGCAACCGCAGCGCAGCGCAA	552			
QY	120	-----LysSerProLysMetSerLysLysAspLysThrValIlePro---	133			
Dd	553	CCTTCGGCTTCCT	612			
QY	134	-----SerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGly	149			
Dd	613	AAAGCAGAACACAGAACACACTAATCTCTCTTTGGGT---GGTAGTCAATGATGTTGGG	669			
QY	150	LeuTy-ProValGlnAsnAsnGluLeuValValThr-GluProAlaGlyPheLeuphePro	169			
Dd	670	ATGTTTCCA-----CCGAAACCGGCG---TTTCTCTCTCCG	702			
QY	170	ValHisAsn-----AspProSerAlaAlaGlnSerAlaPheGlyPhe----	183			
Dd	703	GTCTCCACTGTCGAGGGTTTGAAGGTATCACCCTCTCATCCCAATAGGGTTTCTCTCC	762			
QY	184	GlyAspPheVal-----ValProValValThrGluGluGlyMetAlaPheSerThr	200			



[illegible]

Db QY Db QY Db QY Db QY Db QY Db QY

1550 GCAGCGGGCCACCAACCACGAGGCCGCCTTGTTCGCGCCTTTTGCGCGCGCTGGCGCG 1491  
128 AsplysThr-----ValIleProSerThraAsAlaasPhisCysPheGlyPheValasn 145  
1490 CCGTCGAGCACATCATCTACTCCCCTCTGCGAGCGCG----- 1455  
146 GlnGlutHrGlyLeuTyProValGlnAAsnGluLeuValValThr----- 161  
1454 --GCGGCGCCTTGAAGCGCGCGCGGCATCTCCTCCCGCGAGAACCTCTCTCTCGTCC 1398  
162 -----GluProAlaglyphe----- 166  
1397 TCCTCTCCGACCGCTCTCCGGGTCCAGCAAGTCGGTAGAAACACAGCTGCTGCGCGCTG 1338  
167 -----LeupHeProValHisAsnaeppProSeRAlalaInser 179  
1337 CTGACGTGCGCGCATCGACTTTCTCGCGCGCGCGCGCGAGCCCCAGCTGCC 1278  
180 Ala-----PheGlyPheGlyAspPheValValProValThrGluGluGlyMetAla 197  
1277 GGCTGCCAGCTCTATTACCATGACCCATCCCGCGCGCTGGCACGTGATCATCTCATC 1218  
198 PheSerThrValAsnaSnglyVallAsnLeuGluThrAsnGluAsnPheAsPllysallePro 217  
1217 ACTCTCCGAGTGTCTGCTCTCTCTGCGAGTGG-----CCG 1179  
218 AlaleAsnLeuTyrglyglyaspGly----- 226  
1178 CGAG-CCAGTATATCCGCGCGAGCTCGCGCGCTCTGCGCTGCTGCTGCCACAC 1120  
227 ---AsndyGlyGlyAsnCysPheProPro----- 235  
1119 GCAAACCCGCGCGCATCACCCACCCACCATGACGATCTCACCGAGTCTCGCTCTCGG 1060  
235 ----- 235  
1059 CCTATGCAACGAGCACTAGGCGAGCATCAGCAAGAGACCATGAGCATGATCATC 1000  
236 LeuthrValProleu-----ThrlaeSnGln 244  
999 CTGCTCCAACCCCTCCAAAGTGTTCGACCACTTCATGGACGATGCTGACCGCA 940  
245 SerGlnglulyAszArgvalGlyLeuserygilygluaSpValGlyAspsenVal-Ty 264  
939 CGCGGTGAGCCCGGTGACAGGAGGACGAGAGAGGCGAG--GCTGGGGTACTGCACA 883  
264 rProValargMethrValPheleAsnGluMetPro----- 276  
882 CGGCATCGGCGTCACTGCTGTGTAATAATCCGCGCACCATCATCATCATCATCA 823  
276 ----- 276  
822 TCTTGCTCTCTGTCGANTGATGTCTGCTGCGTGCAGATCTAGTAGCGGGCCATGAT 763  
276 ----- 276  
762 CTTTCCATTACCACTGCTGCTGTCGAGCGCGAGCAACGTCGCTACAGCTGTC 703  
276 ----- 276  
702 ACTCGTGATCATGTTGCAAGCGGCTTTGTCGAGAATTGACGTTGATGGCATCTTC 643  
277 -----IleGluValSerGlyLeu----- 283  
642 TTCTCAAAAAGATGCTTTCAACATCGCCTCTCTCTGCTTGGAATTTCCGATGCATAG 583  
283 ----- 283  
582 ATATGCTAGCTTTTTGGCATCATAAAACTTGCCAGGAGGGGCGCGCGCCGCC 523  
283 ----- 283  
522 CTCGGGACCGGTGGCGGTGAGCCCGNACGAGAGCATGTGTCACCAAGACCCAG 463

```

QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120
Db 199 TTCTACTGTTTCAAAACCGAGCGTCAAGATCTCGC-----CGAGACACCGGACG 249
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
Db 250 CTTTTCAGCAGCCACCGACCGCGCCACCTCCATAGGAGCTGAAGACCAACAGCACATG 309
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValVal 160
Db 310 ACGGCCATGAGCATGATCAA-----TATCCTTGCAGCAACACGAGATT----- 354
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180
Db 355 -----GAT 357
QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200
Db 358 TTGGGGTTTGA-----AGTTGT 375
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219
Db 376 AGCAACTTATCAGCTAATTACTTCCCTTAATGGATCGTCTCATCTCAAAATCCCTTCTTT 435
QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyAsnCys 232
Db 436 TTCCTCGGCCTCTCTTCAAGTGGTGGTGTGAGAACACAAATGGTATGGAGATCTC 495
QY 233 PheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252
Db 496 TTCAAAATGTATGGCCATGAATCTCATCATATCATCATCATCATCATCATCATCATCAT 555
QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267
Db 556 GCTGCATCAGTTTAAACCCATCTCATCAAACTCCAACTCCAAATAGCAACAGAGGG 615
QY 268 ---MetThrValPheIleAsnGluMetProLeuGluValValSerGlyLeuPheAsnVal 286
Db 616 TTTATGACGGTGTATTAACCGAGTTCCTATGGAAGTAACAAAGAGGACATAGACATG 675
QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306
Db 676 AAAACAATGTCGTGTATGATCGGTGTATCTTCACTTCTCTGCTCTCTCTCTCTCT 735
QY 307 AspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325
Db 736 GATGAGTTTGGTTTCTTGTGATGATCTTTTACACATGACAAACTTATTTCTCTGGTA 792

```

## RESULT 7

```

US-10-225-066A-165
; Sequence 165, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBEL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: FILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: FINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: W810038-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18

```

```

; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 165
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-225-066A-165

```

## Alignment Scores:

```

Pred. No.: 3,84e-40 Length: 807
Score: 394.50 Matches: 105
Percent Similarity: 47.83% Conservative: 38
Best Local Similarity: 35.12% Mismatches: 101
Query Match: 22.53% Indels: 55
DB: 13 Gaps: 9

```

US-09-787-737-2 (1-325) x US-10-225-066A-165 (1-807)

```

QY 41 SerSerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60
Db 19 TCACATAGCCCATCTCTCCACTTCCACCGAACAGTCCGGGCACGTTGGTCACTAAACCG 78
QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProArgGlu 80
Db 79 GAGCAAAATCTTGATCTCGAATCCATCTTCAACAGTGGTACTGTATAACCCACCAAGAT 138
QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnVal 100
Db 139 GAAACGGTGGAGATAAGAAAGATGCTTGAGAAATTCGTGCTGTGGGAGACCAACGTC 198
QY 101 PheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLys 120
Db 199 TTCTACTGTTTCAAAACCGAGCGTCAAGATCTCGC-----CGAGACACCGGACG 249
QY 121 SerProLysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCys 140
Db 250 CTTTTCAGCAGCCACCGACCGCGCCCTCTCATAGGAGCTGAAGACCAACAGCATG 309
QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrProValGlnAsnGluLeuValVal 160
Db 310 ACGGCCATGAGCATGATCAA-----TATCCTTGCAGCAACACGAGATT----- 354
QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180
Db 355 -----GAT 357
QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200
Db 358 TTGGGGTTTGA-----AGTTGT 375
QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnPheAspLysIleProAlaIle 219
Db 376 AGCAACTTATCAGCTAATTACTTCCCTTAATGGATCGTCTCATCTCAAAATCCCTTCTTT 435
QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyAsnCys 232
Db 436 TTCCTCGGCCTCTCTTCAAGTGGTGGTGTGAGAACACAAATGGTATGGAGATCTC 495
QY 233 PheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252
Db 496 TTCAAAATGTATGGCCATGAATCTCATCATATCATCATCATCATCATCATCATCATCAT 555
QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrProValArg--- 267
Db 556 GCTGCATCAGTTTAAACCCATCTGATCAAACTCCAACTCCAAATAGCAACAGAGGG 615

```

QY 268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286  
 Db 616 TTTATGACGGTGTATTAAACGGAGTTCCTATGAAAGTAAACAAAGGAGCAATAGACATG 675  
 QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306  
 Db 676 AAACAAATGTTCCGTGATGATTCGGTGTACTTCACTTCTCTCTCTCTCTCTCTCTCCACT 735  
 QY 307 AspGluPheGlyValThrTyrrGlnProLeuGlnAsnGlyAlaIleTyrrTyrrLeuIle 325  
 Db 736 GATGAGTTTGGTTTCTTGTATGATCAATCTTATACATGACAACTATATTTCTTGTA 792

RESULT 8

US-10-374-780A-2665  
 ; Sequence 2665, Application US/10374780A  
 ; Publication No. US20040019927A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sherman, Bradley K  
 ; APPLICANT: Riechmann, Jose Luis  
 ; APPLICANT: Jiang, Cai-Zhong  
 ; APPLICANT: Heard, Jacqueline E  
 ; APPLICANT: Haake, Volker  
 ; APPLICANT: Creelman, Robert A  
 ; APPLICANT: Ratcliffe, Oliver  
 ; APPLICANT: Adam, Luc J  
 ; APPLICANT: Reuber, T. Lynne  
 ; APPLICANT: Reddie, James  
 ; APPLICANT: Broun, Pierre E  
 ; APPLICANT: Pilgrim, Marsha L  
 ; APPLICANT: Dubell III, Arnold T  
 ; APPLICANT: Pineda, Omaira  
 ; APPLICANT: Yu, Guo-Liang  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

FILE REFERENCE: MBI-0047 CLP  
 CURRENT APPLICATION NUMBER: US/10/374,780A  
 CURRENT FILING DATE: 2003-02-25  
 PRIOR APPLICATION NUMBER: 09/837,944  
 PRIOR FILING DATE: 2001-04-18  
 PRIOR APPLICATION NUMBER: 60/310,847  
 PRIOR FILING DATE: 2001-08-09  
 PRIOR APPLICATION NUMBER: 09/934,455  
 PRIOR FILING DATE: 2001-08-22  
 PRIOR APPLICATION NUMBER: 60/336,049  
 PRIOR FILING DATE: 2001-11-19  
 PRIOR APPLICATION NUMBER: 60/338,692  
 PRIOR FILING DATE: 2001-12-11  
 PRIOR APPLICATION NUMBER: 10/171,468  
 PRIOR FILING DATE: 2002-06-14  
 PRIOR APPLICATION NUMBER: 10/225,066  
 PRIOR FILING DATE: 2002-08-09  
 PRIOR APPLICATION NUMBER: 10/225,067  
 PRIOR FILING DATE: 2002-08-09  
 PRIOR APPLICATION NUMBER: 10/225,068  
 PRIOR FILING DATE: 2002-08-09  
 NUMBER OF SEQ ID NOS: 2906  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 2665  
 LENGTH: 807  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 FEATURE:  
 ; OTHER INFORMATION: GI:586  
 US-10-374-780A-2665

Alignment Scores:  
 Pred. No.: 3,846-40  
 Score: 394.50  
 Conservativity: 47.8%  
 Best local Similarity: 35.12%  
 Indels: 55  
 Query Match: 16  
 DB: 9

US-09-787-737-2 (1-325) x US-10-374-780A-2665 (1-807)

QY 41 SerSerPheSerSerAspArgIleProAspProLysProArgTrpAsnProLysPro 60  
 Db 19 TCATATAGCCCATCTCTCCATCCACCGAACAGTCCGGGCGACGTGGTCACTAAACCG 78  
 QY 61 GluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProArgGlu 80  
 Db 79 GAGCAAACTTGTATCTCGAATCCATCTTCAACAGTGGTACTGTGTAAACCCCAAAAGAT 138  
 QY 81 GluIleGlnArgIleArgIleArgLeuGlnGluTyrrGlyGlnIleGlyAspAlaAsnVal 100  
 Db 139 GAAACGGTGAAGATAGAAAGATGCTTACAAATTCGTGCTGTGGGAGACGCAACAGTC 198  
 QY 101 PheTyrrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLys 120  
 Db 199 TTCTACTGGTTTCAAAACCGAGCGGTCAAGATCTCGC-----CGAGACACCGGAG 249  
 QY 121 SerProLysMetSerLysAspLysThrValIleProSerThrAspAlaAspHisCys 140  
 Db 250 CTTTTCAGCAGCCACCGCAGCGCCACCTCCATAGAGCTGAAGCACCACGACACATG 309  
 QY 141 PheGlyPheValAsnGlnGluThrGlyLeuTyrrProValGlnAsnAsnGluLeuVal 160  
 Db 310 ACGGCCATGAGCATGCATCAA-----TATCCTTCAGCAACAACAGAGATT----- 354  
 QY 161 ThrGluProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAla 180  
 Db 355 -----GAT 357  
 QY 181 PheGlyPheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThr 200  
 Db 358 TTGGGGTGTGGA-----AGTTGT 375  
 QY 201 ValAsnAsnGlyValAsnLeuGluThrAsn---GluAsnAspAspLysIleProAlaIle 219  
 Db 376 AGCAACTTATCAGCTAACTTACTTCTTAATGGATCGTGTCTCATCTCAAAATCCCTTCTTT 435  
 QY 220 AsnLeu-----TyrGlyGly-----AspGlyAsnGlyGlyGlyAsnCys 232  
 Db 436 TTCTTCGGCTCTCTTCTTCAAGTGGTGTGAGAACAAACAATGGTATGGAGATCTC 495  
 QY 233 PheProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGly 252  
 Db 496 TTCAAAATGTATGGCCATGAATCTGATCATATATCATCAGCAGCAGCATCATCAAT 555  
 QY 253 -----LeuSerGlyGlyGluAspValGlyAspAsnValTyrrProValArg--- 267  
 Db 556 GCTGCATCAGTTTTTAAACCCATCTGATCAAACTCCCACTCCCAATACGAAAGAGG 615  
 QY 268 ---MetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286  
 Db 616 TTTATGACGGTGTATTAAACGGAGTTCCTATGAAAGTAAACAAAGGAGCAATAGACATG 675  
 QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGlyGlnProIleLeuThr 306  
 Db 676 AAACAAATGTTCCGTGATGATTCGGTGTACTTCACTTCTCTCTCTCTCTCTCTCCACT 735  
 QY 307 AspGluPheGlyValThrTyrrGlnProLeuGlnAsnGlyAlaIleTyrrTyrrLeuIle 325  
 Db 736 GATGAGTTTGGTTTCTTGTATGATCAATCTTATACATGACAACTATATTTCTTGTA 792

RESULT 9  
 US-10-437-963-59649  
 ; Sequence 59649, Application US/10437963  
 ; Publication No. US20040123343A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 59649  
 ; LENGTH: 1524  
 ; TYPE: DNA  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_61253C.1  
 US-10-437-963-59649

Alignment Scores:  
 Pred. No.: 4,76e-36 Length: 1524  
 Score: 366.00 Matches: 128  
 Percent Similarity: 35.61% Conservative: 44  
 Best Local Similarity: 26.50% Mismatches: 111  
 Query Match: 20.90% Indels: 200  
 DB: 17 Gaps: 16

US-09-787-737-2 (1-325) x US-10-437-963-59649 (1-1524)  
 QY 41 SerSerSerPheSerSer-----AspArgIleProAspProLysProArgTTPAsn 57  
 Db TCGTGGCCCTTCTCGTCAGTGGCGAGGAGAGGGTTCCGGACCCGAGCGCGGTGGAAC 135  
 QY 58 ProLysProGluGlnIleArgIleGluSerIlePheAsnSerGlyThrIleAsnPro 77  
 Db CGCGCGCGGACGATCCGGATCCGATCGGAGGCGGATCTCACTCCGGCATGTCAACCG 195  
 QY 78 ProArgGluGluIleGlnArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAsp 97  
 Db CGCGCGCGGAGATCCCGGATCCGATCGATCGAGTGCAGGAGTACGGCAGTCCGGCGAC 255  
 QY 98 AlaAsnValPheTyrPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg--- 116  
 Db GCCAACGCTTCTACTGTGTTCCAGAACCGCAAGTCCCGCTCCAGAACAGTGGCGTCC 315  
 QY 116 ----- 116  
 Db GCGCGGACAGCGCGCGGGCTCGGCTCGCGCGCAACCGCGCTCGCGCGCGCGCG 375  
 QY 117 ValHisHisLys----- 120  
 Db GCGCAGCGGAGGCGGTGGCGCGCTGTTACGCGCGCGCCACCAATCTCCCGCGCGCC 435  
 QY 121 ----- 121  
 Db CAGCGCGTGCAGCGCGAGCAGCAGCTGTCTCGCTGTGGCGCGCTACTCTGCTGTCG 495  
 QY 124 MetSerLysLysAspLysThr----- 130  
 Db TCTTCTCTCTCCACCGCTTCTCGGTCCAGCAAGCTTCGAGGGCTACGTTCGACGCGCAG 555  
 QY 131 -----ValIleProSerThrAspAlaAspHis--- 139  
 Db GCGATGTCGCTGACGAGCGCCATGAGCCTGCTCTCGCGCTCGCGCGCGGTGCCACCG 615  
 QY 139 ----- 139  
 Db CAGATGCTCTATCAGGCGCCAGCCTGGAGTCCGCGCGCGCGCTGCTCCCAAGTGCAC 675  
 QY 140 -----CysPheGly 142  
 Db GGCATGCTGCCACAGCAGCGCGGTCTTCTCGAGTGGCGCGAGCCCTGCTGCTGTCG 735  
 QY 143 PheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValThrGlu 162  
 Db GCGCGTGCAGCTCGCGCGCGCCATCTTGGCGGCGCAGTACATGACCTGCGCGTGGCGGT 795

QY 163 Pro-----AlaGlyPheLeuPheProValHisAsnAsp--- 173  
 Db CCGCAGCCACCGTCTGCGCGCGCGCGCGCGGTCTGCGGGCTCTCAACACGCGTG 855  
 QY 174 -----ProSerAlaAlaGlnSerAlaPheGlyPheGlyAsp 185  
 Db CAAGCGCCAAACACACCGGCCACAGAGTGGCGCTGGAGCGCC---GGCTCGGCCAG 912  
 QY 186 PheValValProValValThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyVal 205  
 Db CACTGGTGGCGCTCCGCGCATCAGCTCGCGCTCGCAAGAGCAGCGCGCGGTGATCGCC 972  
 QY 206 AsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsn-----LeuTyr 222  
 Db ACCGTGTCTAGGCGGAGGAGGCGCAGCAGTCCAGCCACAGCAGCGGTCTGTCTACAG 1032  
 QY 223 GlyGlyAspGly-----AsnGlyGlyGlyAsn 231  
 Db TACGGCTTTGGCATCACCACCGCGCAAGTGCACGTGCAGCTTACCTCTCTCGGTGTGGC 1092  
 QY 232 CysPheProProLeu----- 236  
 Db GTTCTGCT 1152  
 QY 237 -----ThrValProLeuThr----- 241  
 Db GTGGCGCGCCACCGCTAGCTGACTGATTTGTTGCTGCAAGTGTATATCTGCTGGCGCGTC 1212  
 QY 242 ---IleAsnGlnSerGlnGlnLysArgAspValGlyLeuSerGlyGly----- 256  
 Db GCTAACATCATGTTTCAAGGCTCTCGCGGATTTTCGGGCTCTGTCGCGCGCTCTCTCTCT 1272  
 QY 257 -----GluAspValGlyAspAsnValValValValValValValValValValVal 266  
 Db GCGGAGCG 1332  
 QY 267 ArgMetThrValPheIleAsnGluMetProIleGluValValSerGlyLeuPheAsnVal 286  
 Db TGGCTGAGCGTGGCG 1392  
 QY 287 LysAlaAlaPheGlyAsnAspAlaValLeuIleAsnSerPheGly-----Gln 302  
 Db AGG---CACTACGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1449  
 QY 303 ProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyr 322  
 Db CTTGTGCGCGTGCAGCAGTGGCGGTCCAGCGCTCAGCGCGCTCCAGCGCGCGCGCGCTCTAC 1509  
 QY 323 TyrLeuIle 325  
 Db ATTGTTGTC 1518  
 RESULT 10  
 US-10-425-114-12809  
 ; Sequence 12809, Application US/10425114  
 ; Publication No. US2004003488A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack B.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 12809  
 ; LENGTH: 1215  
 ; TYPE: DNA



```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701212767_FLI
US-10-425-114-12809

Alignment Scores:          9,89e-35      Length: 1215
Pred. No.:                354.50         Matches: 100
Score:                     43.98%        Conservative: 46
Percent Similarity:        30.12%        Mismatches: 105
Best Local Similarity:     20.25%        Indels: 81
Query Match:              13             Gaps: 10
DB:

US-09-787-737-2 (1-325) x US-10-425-114-12809 (1-1215)

Qy 24 GlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeuSerSer 43
Db 90 CAGCATCAAGTGGAGCAAGAGTCCAAAGGCAA-----122
Qy 44 PheSerSerAspArgIleProAspProLysProArgTyrAsnProLysProGluGlnIle 63
Db 123 ---GGCACCAGAGAGAGTGAAGCGTTCGGTCAAGGTGACTCCCAAGCCAGAACAAATC 179
Qy 64 ArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluLeuGln 83
Db 180 CTCATCTCAGTCCATCTTCAACAGTGGCATGTTAAACCTCCAAAGGAGCAACCGTTC 239
Qy 84 ArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrp 103
Db 240 AGATAAGGAAGCTTCTCGAGAAATTCGGCGCGCGTCCGCGACGCCCAACGTTCTTACTGG 299
Qy 104 PheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLys 123
Db 300 TTCCAGAACCGCGCTCCAGATCCCGCCCGCCAGCGCATGATGCAACAGCAGGCT 359
Qy 124 MetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPhe 143
Db 360 ACTGCCACCGCCACCGCCACCGTC-----ACCACCTTTGATCAC-----398
Qy 144 ValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValValThrGluPro 163
Db 399 -----CCTCAGCCTCAGACTCAGACTCTTGTAAAT-----428
Qy 164 AlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPhe 183
Db 429 GTTGGGTGGTGCATTCCTCA-----CATGATCACCACACC-----TTG 464
Qy 184 GlyAspPheValValProValProValThrGluGluGlyMetAlaPheSerThrValAsnAsn 203
Db 465 GGAACCTT-----GTTGTGCAACTGAAAGTAGTGTCTACTTCAACCATGGGTTT 515
Qy 204 GlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyrGly 223
Db 516 GGTGTGTTCTTCTCAATCTCTTATGGTTTCTTGGTTCCTCTCTTCTTCTTCTTCTTGT 575
Qy 224 GlyAspGlyAsnGlyGlyGlyAsn-----231
Db 576 GGTGGTGTGTAAATGTGGTGGCCACAAACAGGCATGGATGTTTCTTCTCATCGGTTTCT 635
Qy 232 -----CysPhe 233
Db 636 TCTCATCAATAGGGTTTCCCTGATCACCACCACTTCCACTGCTGTTTCATGAGTTTGTAC 695
Qy 234 ProProLeuThrValProIleuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeu 253
Db 696 CCTCCTCTGTATCCAAATTTGACCTACACGGCTGGATAT-----734
Qy 254 SerGlyGlyGluAspValGlyAspAsnValTyrProValArgMetThrValPheIleAsn 273
Db 735 ---GGGGGCCCTAATATTTCAGGAATCATT-----ACAGTGTATTCATCAAT 776
Qy 274 GluMetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAsp 293

```

```

777 GGAATTGCAACTGAACATTCCTCAAGGGCCCAATAGACCTCAAAACCGTGTGAGAGAGAT 836
294 AlaValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyr 313
837 GTAATGTTAGTTCATTCCTCTGGAGTTCCTCAATTCCTCAATTCCTCAATTCCTCAATTC 896
314 GlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325
897 CATCACTTCACATGCGATGCTACTTCTCGTA 932

RESULT 11
US-10-424-599-97798
Sequence 97798, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 97798
LENGTH: 1218
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_59325C.1
US-10-424-599-97798

Alignment Scores:          9,93e-35      Length: 1218
Pred. No.:                354.50         Matches: 100
Score:                     43.98%        Conservative: 46
Percent Similarity:        30.12%        Mismatches: 105
Best Local Similarity:     20.25%        Indels: 81
Query Match:              13             Gaps: 10
DB:

US-09-787-737-2 (1-325) x US-10-424-599-97798 (1-1218)

Qy 24 GlnHisGluLeuAspThrProSerTyrMetHisTyrSerAsnCysAsnLeuSerSer 43
Db 90 CAGCATCAAGTGGAGCAAGAGTCCAAAGGCAA-----122
Qy 44 PheSerSerAspArgIleProAspProLysProArgTyrAsnProLysProGluGlnIle 63
Db 123 ---GGCACCAGAGAGAGTGAAGCGTTCGGTCAAGGTGAGTCTCCAAAGCCAGAACAAATC 179
Qy 64 ArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluLeuGln 83
Db 180 CTCATCTCAGTCCATCTTCAACAGTGGCATGTTAAACCTCCAAAGGAGCAACCGTTC 239
Qy 84 ArgIleArgIleArgLeuGlnGluTyrGlyGlnIleGlyAspAlaAsnValPheTyrTrp 103
Db 240 AGATAAGGAAGCTTCTCGAGAAATTCGGCGCGCGTCCGCGACGCCCAACGTTCTTACTGG 299
Qy 104 PheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisLysSerProLys 123
Db 300 TTCCAGAACCGCGCTCCAGATCCCGCCCGCCAGCGCATGATGCAACAGCAGGCT 359
Qy 124 MetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGlyPhe 143
Db 360 ACTGCCACCGCCACCGCCACCGTC-----ACCACCTTTGATCAC-----398
Qy 144 ValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValValThrGluPro 163
Db 399 -----CCTCAGCCTCAGACTCAGACTCTTGTAAAT-----428
Qy 164 AlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGlyPhe 183

```

Db 429 GTTGGTGGTCAATTCCA---CATGATCACCACACC-----TTG 464  
Qy 184 GlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThrValAsnAsn 203  
Db 465 GGAACACCTT-----GTTGTTGCAACTGAAAGTAGTGTACTTCAACCATGGGTTT 515  
Qy 204 GlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuThrGly 223  
Db 516 GGTGTTCTTCTCAATCTTCTATTAGGTTTCTTGGTTCTTCTTCTTCTTCTTCTTCTTGT 575  
Qy 224 GlyAspGlyAsnGlyGlyGlyAsn----- 231  
Db 576 GGTGGTGGTGAATTGGTGGCCCAACAACAGGCATGATGTTCTTCTCATCGGTTTCT 635  
Qy 232 -----CysPhe 233  
Db 636 TCTCATCAAAATGGTTTCCCTGTATCACCACACACTTCACTGCTTTCATCAGCTTGTATC 695  
Qy 234 ProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLysArgAspValGlyLeu 253  
Db 696 CTTCTCTTGTATCAATTTGACCTTACCAGGCTGGATAT----- 734  
Qy 254 SerGlyGlyGluAspValGlyAspAsnValThrProValArgMetThrValPheIleAsn 273  
Db 735 ---GGGGGCCCTAAATATTTCAGGATTCATT-----ACAGTGTATTATCAAT 776  
Qy 274 GluMetProIleGluValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAsp 293  
Db 777 GGAATTTGCACTGAACCTTCAAAAGGGCCCAATAGACTCAAAACCGTTTGGAGAGAT 836  
Qy 294 AlaValLeuIleAsnSerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyr 313  
Db 837 GTAAGTTAGTTTCATCTCTGGAGTTCCAAATCCCAACCAATGAATTTGGGTTCTTGATG 896  
Qy 314 GlnProLeuGlnAsnGlyAlaIleTyrTyrLeuIle 325  
Db 897 CATACCTTCAACATGGCGATAGTACTTCTGGTA 932

RESULT 12

US-09-934-455-489  
; Sequence 489, Application US/09934455  
; Publication No. US20030121070A1  
; GENERAL INFORMATION:  
; APPLICANT: Adam, Luc  
; APPLICANT: Creelman, Robert  
; APPLICANT: Dubell, Arnold  
; APPLICANT: Heard, Jacqueline  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Keddie, James  
; APPLICANT: Pilgrim, Marsha  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Reuber, Lynne  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: Pineda, Omalra  
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
; FILE REFERENCE: MBI-0025  
; CURRENT APPLICATION NUMBER: US/09/934,455  
; PRIOR APPLICATION NUMBER: 2001-08-22  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: MBI-0022  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: MBI-0023  
; PRIOR FILING DATE: 2001-04-17  
; NUMBER OF SEQ ID NOS: 516  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 489  
; LENGTH: 994  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: (1)..(711)  
; OTHER INFORMATION: G2547  
US-09-934-455-489

Alignment Scores:

Pred. No.: 2.89e-33 Length: 994  
Score: 342.00 Matches: 87  
Percent Similarity: 44.22% Conservative: 47  
Best Local Similarity: 28.71% Mismatches: 95  
Query Match: 19.53% Indels: 75  
DB: 10 Gaps: 4

US-09-787-737-2 (1-325) x US-09-934-455-489 (1-994)

Qy 23 HisGlnHisGluIleAspThrProSerTyrMetHisTyrSerAsnCysAsnLeuSerSer 42  
Db 11 AACAAACACCATACATCCCAACCCGATAGTCGTCACCCCATCC-TCCGCTCCGGT 69  
Qy 43 SerPheSerSerAspArgIleProAspProLysProArgTyrAsnProLysProGluGln 62  
Db 70 TCCACCTCAGCAACACCGTT-----CGTCCCGATGTCTCACTAAACCGAACA 120  
Qy 63 IleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnProProArgGluGluIle 82  
Db 121 ATACTCATCTTGAAGTCGATCTTCCACAGTGTATGTGTAACTCCCAAGAGAGAG 180  
Qy 83 GlnArgIleArgIleArgLeuGlnGluThrGlyGlnIleGlyAspAlaAsnValPheTyr 102  
Db 181 GTAAGGATAAGAAAGATGCTCGAGAAATTTGGCGCGGTGGGAGATGCCAATGTCTTCTAT 240  
Qy 103 TrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArgValHisHisLysSerPro 122  
Db 241 TGGTTTCAAAACCGCGGTCAAGGTCCCGTCGGACACAGCGA----- 282  
Qy 123 LysMetSerLysLysAspLysThrValIleProSerThrAspAlaAspHisCysPheGly 142  
Db 282 ----- 282  
Qy 143 PheValAsnGlnGluThrGlyLeuTyrProValGlnAsnAsnGluLeuValValThrGlu 162  
Db 282 ----- 282  
Qy 163 ProAlaGlyPheLeuPheProValHisAsnAspProSerAlaAlaGlnSerAlaPheGly 182  
Db 283 -----CAGCTACAGGCTGCAGCTGCA 303  
Qy 183 PheGlyAspPheValValProValValThrGluGluGlyMetAlaPheSerThrValAsn 202  
Db 304 GCAGCGGAC-----GCAACCCACCAACACTTCTAGTCTTCTTCTTATGGT 351  
Qy 203 AsnGlyValAsnLeuGluThrAsnGluAsnPheAspLysIleProAlaIleAsnLeuTyr 222  
Db 352 GGTGGATGTGATTAATCAAGCAATAGTGGCATGGAGATCTCTTAAACAATGTCT- 405  
Qy 223 GlyGlyAspGlyAsnGlyGlyAsnCysPheProLeuThrValProLeuThrIle 242  
Db 406 -----GGCCAAATGAGTACCATTGAGTGGTCCCATCTGATCAAAACTCCAAATTTCAA 507  
Qy 243 AsnGlnSerGlnGluLysArgAspValGlyLeuSerGlyGlyGluValValSerGly 282  
Db 448 AATCATAGCTCAAAATGTCATCGATTTTGTGCCCATCTGATCAAAACTCCAAATTTCAA 507  
Qy 263 ValTyrProValArgMetThrValPheIleAsnGluMetProIleGluValValSerGly 282  
Db 508 TACCACCAAGGGCTTATACCGTGTATTAAACGGAGTTCCGACAGAGTCCGACGAGGA 567  
Qy 283 LeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValIleAsnSerPheGlyGln 302  
Db 568 GGAATAGACATCAAAAGCAACGTTTGGAGAGATTTGGTTTGGTGCATCTCCAGTGT 627  
Qy 303 ProfileLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsnGlyAlaIleTyr 322  
Db 628 CTTCTTCTACTGATGATTTGGTTTGTGATGATAGTCTTACACATGGTGGTCAAGCTTAT 687

```

QY 323 TyrLeuIle 325
Db 688 TTCCTGGTT 696

RESULT 13
US-10-260-238-1567
; Sequence 1567, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goffi, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Yong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; -SEQ ID NO 1567
; -LENGTH: 837
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-1567

Alignment Scores:
Pred. No.: 7,13e-33 Length: 837
Score: 338.00 Matches: 89
Percent Similarity: 46.34% Conservative: 44
Mismatch: 31.01% Mismatches: 92
Query Match: 19.30% Indels: 62
DB: 9 Gaps: 16

US-09-787-737-2 (1-325) x US-10-260-238-1567 (1-837)
QY 53 LysProArgTrpAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSer 72
Db 58 AGTTCGGTGGAGCGGAGCGGAGGAGATCTATCTCTGGAGTCCATCTTCAACAGC 117
QY 73 GlyThrIleAsnProArgGluGluIleGlnArgIleArgLeuGlnGluTyr 92
Db 118 GGCATGTTGAACCGCGCCAGAGGAGAGCCGTCGCGATCCGCAAGCTCTCGAGCGTTC 177
QY 93 GlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLys 112
Db 178 GCGCCCTGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 237
QY 113 HisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIle 132
Db 238 CGCGCCGCGCGC----- 249
QY 133 ProSerThrAspAlaAsnHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrPro 152
Db 250 ---CAGCTGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 306
QY 153 ValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHis--- 171
Db 307 -----ACTGTTCTCGGTGGGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCG 342
QY 172 AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValProVal 191

```

```

Db 343 GGCTCGCGCGCTTCGTCTCGGATGTTTCGGCGACGCGC----- 381
QY 192 ThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGlu 211
Db 382 ---GCCCGCGGTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 423
QY 212 AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAsp-----GlyAsnGly 228
Db 424 ---TCGCGCGCGTTCGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 474
QY 229 GlyGlyAsnCysPheProProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys 248
Db 475 GCGCAGCAGCTGTTTC-----GCCATCTCGAGCGCAGATGGGGTAC 513
QY 249 ArgAspValGly-----LeuSerGlyGlyGluAspValGlyAspAsn 262
Db 514 ATGACGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 573
QY 263 ValTyr-----ProValArgMetThrValPheIleAsnGluMetProIleGlu 278
Db 574 CTCTACTACTCTCTCAACCTCGACGATGATGATGATGATGATGATGATGATGATGATGATG 633
QY 279 ValValSerGlyLeuPheAsnValLysAlaPheGlyAsnAspAlaValLeuIleAsn 298
Db 634 GTGCCAGAGGAGCGGATCGATCTGAGATCAATGTTTGGGAGGAGCTGATGCTGTGTCAT 693
QY 299 SerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsn 318
Db 694 TCACGGGTGCTCTCTCTCCAGCCACGAGTACGGCATCTCTCTCTCTCTCTCTCTCTCTCT 753
QY 319 GlyAlaIleTyrTyrLeuIle 325
Db 754 GCGCAGAGCTACTTCTCTGTT 774

RESULT 14
US-10-437-963-43264
; Sequence 43264, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 43264
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46438C.1
US-10-437-963-43264

Alignment Scores:
Pred. No.: 7,13e-33 Length: 837
Score: 338.00 Matches: 89
Percent Similarity: 46.34% Conservative: 44
Mismatch: 31.01% Mismatches: 92
Query Match: 19.30% Indels: 62
DB: 9 Gaps: 17

US-09-787-737-2 (1-325) x US-10-437-963-43264 (1-837)
QY 53 LysProArgTrpAsnProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSer 72

```

```

Db 58 AGTTCGGGGTGGACCGCCAAACCGGACAGATCATCTCTGGAGTCCATCTTCAACAGC 117
Qy GlyThrIleAsnProProArgGluGluIleGlnArgIleArgLeuGlnGluTyr 92
Db 118 GGCATGTGTGAACCGCCCAAGGACGACCGTCCGCATCCGCAAGCTGCTCGAGCGTTC 177
Qy GlyGlnIleGlyAspAlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLys 112
Db 178 GGCAGCGCTGCGACGACGCAACGCTTCTACTGTGTTCCAGAACCGCGCTCGCGTCCCGC 237
Qy HisLysLeuArgValHisHisLysSerProLysMetSerLysLysAspLysThrValIle 132
Db 238 CGCGGCAGCGC-----249
Qy ProSerThrAspAlaAspHisCysPheGlyPheValAsnGlnGluThrGlyLeuTyrPro 152
Db 250 ---CAGTCGACGCGCAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
Qy ValGlnAsnAsnGluLeuValValThrGluProAlaGlyPheLeuPheProValHis--- 171
Db 307 -----ACTGCTTCGTCGGTGGCTCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCG 342
Qy AsnAspProSerAlaAlaGlnSerAlaPheGlyPheGlyAspPheValProValVal 191
Db 343 GAGTCGCCCGCTTCGTCGTCGCGGATGTTCCGCGACGCG-----381
Qy ThrGluGluGlyMetAlaPheSerThrValAsnAsnGlyValAsnLeuGluThrAsnGlu 211
Db 382 ---GCCGCGCGGTACAGCTCTCTGTCGTCGTCCTCATGTCGCGCGTCC-----423
Qy AsnPheAspLysIleProAlaIleAsnLeuTyrGlyGlyAsp-----GlyAsnGly 228
Db 424 -----TCGCCGCGCGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 474
Qy GlyGlyAsnCysPheProLeuThrValProLeuThrIleAsnGlnSerGlnGluLys 248
Db 475 GCGCAGCAGCTGTTTC-----GCCATCTCGAGGCGAGATGGGTAC 513
Qy ArgAspValGly-----LeuSerGlyGlyGluAspValGlyAspAsn 262
Db 514 ATGACGCGCGCGCGCGTCTGTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 573
Qy ValTyr-----ProValArgMetThrValPheIleAsnGluMetProIleGlu 278
Db 574 CTCCTACTCTGTCGTCACCTCGGACGATGACGCTGTTTCATCAACGAGTGGCGACGAG 633
Qy ValValSerGlyLeuPheAsnValLysAlaAlaPheGlyAsnAspAlaValIleAsn 298
Db 634 GTGCAGGGGACCGATCGATCGATCAATGTTTGGCGAGGACGTCGTCGTCGTCGTCGTCG 693
Qy SerPheGlyGlnProIleLeuThrAspGluPheGlyValThrTyrGlnProLeuGlnAsn 318
Db 694 TCAACGGGTGCTCTCTTCAGCAACAGATGACGATCGGATCTCTCTCTCTCTCTCTCTCT 753
Qy GlyAlaIleTyrLeuIle 325
Db 754 GCGGAGAGTACTTCTCTGTT 774

```

## RESULT 15

```

US-10-437-963-59651
; Sequence 59651, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

```

```

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 59651
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61255C.1
US-10-437-963-59651

```

## Alignment Scores:

```

Pred. No.: 1,02e-31 Length: 921
Score: 329.50 Matches: 61
Percent Similarity: 91.14% Conservative: 11
Best Local Similarity: 77.22% Mismatches: 4
Query Match: 18.82% Indels: 3
DB: 17 Gaps: 1

```

```

US-09-787-737-2 (1-325) x US-10-437-963-59651 (1-921)

```

```

Qy 41 SerSerSerPheSerSer-----AspArgIleProAspProLysProArgTrpAsn 57
Db 355 TCCTCGCCCTTCTCGTCAGTGGCGGAGAGAGGTTCGCGACCCGAAAGCGCGCGGAAC 414
Qy 58 ProLysProGluGlnIleArgIleLeuGluSerIlePheAsnSerGlyThrIleAsnPro 77
Db 415 CGCGCGCGGACGACATCCGGATCTCTGGAGGCCATCTTCACTCCGCGCATGTGTCAACCG 474
Qy 78 ProArgGluGluIleGlnArgIleArgLeuGlnGluTyrGlyGlnIleGlyAsp 97
Db 475 CGCGCGGACGAGATCCCGCGCATCCGCGATCGCATGCGAGTACGCGCAGTCCGCGGAC 534
Qy 98 AlaAsnValPheTyrTrpPheGlnAsnArgLysSerArgAlaLysHisLysLeuArg 116
Db 535 GCCAACGCTTTCTACTGTTCCAGAACCGCAAGTCCCGCTCCAGAACAGAGTGGCGC 591

```

```

Search completed: September 10, 2004, 12:59:25
Job time : 581 secs

```

**This Page Blank (uspto)**